

08/03/00
1c853 U.S. PTO

08-04-00

A
Box 500

Please type a plus sign (+) inside this box → ☐

PTO/SB/05 (4/98)
Approved for use through 09/30/2000. OMB 0651-0032
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 C.F.R. § 1.53(b))

Attorney Docket No. 12/211
First Inventor or Application Identifier Konopitzky, Renate
Title Tumor-associated antigen (R11)
Express Mail Label No. EL580603893US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. ☒ * Fee Transmittal Form (e.g., PTO/SB/17)
(Submit an original and a duplicate for fee processing)
2. ☒ Specification [Total Pages 44]
(preferred arrangement set forth below)
 - Descriptive title of the Invention
 - Cross References to Related Applications
 - Statement Regarding Fed sponsored R & D
 - Reference to Microfiche Appendix
 - Background of the Invention
 - Brief Summary of the Invention
 - Brief Description of the Drawings (if filed)
 - Detailed Description
 - Claim(s)
 - Abstract of the Disclosure
3. ☒ Drawing(s) (35 U.S.C. 113) [Total Sheets 6]
4. Oath or Declaration (unexecuted) [Total Pages 3]
 - a. ☐ Newly executed (original or copy)
 - b. ☐ Copy from a prior application (37 C.F.R. § 1.63(d))
(for continuation/divisional with Box 16 completed)
 - i. ☐ DELETION OF INVENTOR(S)
Signed statement attached deleting
inventor(s) named in the prior application,
see 37 C.F.R. §§ 1.63(d)(2) and 1.33(b).

* NOTE FOR ITEMS 1 & 13: IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY
FEES, A SMALL ENTITY STATEMENT IS REQUIRED (37 C.F.R. § 1.27), EXCEPT
IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C.F.R. § 1.28).

ADDRESS TO:

Assistant Commissioner for Patents
Box Patent Application
Washington, DC 20231

5. ☐ Microfiche Computer Program (Appendix)
6. Nucleotide and/or Amino Acid Sequence Submission
(if applicable, all necessary)
 - a. ☒ Computer Readable Copy
 - b. ☒ Paper Copy (identical to computer copy)
 - c. ☒ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

7. ☐ Assignment Papers (cover sheet & document(s))
8. ☐ 37 C.F.R. § 3.73(b) Statement of Power of Attorney
(when there is an assignee)
9. ☐ English Translation Document (if applicable)
10. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
11. ☐ Preliminary Amendment
12. ☒ Return Receipt Postcard (MPEP 503)
(Should be specifically itemized)
13. ☐ * Small Entity Statement filed in prior application,
Statement(s) Status still proper and desired
(PTO/SB/09-12)
14. ☐ Certified Copy of Priority Document(s)
(if foreign priority is claimed)
15. ☐ Other:

16. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment:
☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No: _____

Prior application information: Examiner: _____ Group / Art Unit: _____
For CONTINUATION or DIVISIONAL APPS only: The entire disclosure of the prior application, from which an oath or declaration is supplied
under Box 4b, is considered a part of the disclosure of the accompanying continuation or divisional application and is hereby incorporated by
reference. The incorporation can only be relied upon when a portion has been inadvertently omitted from the submitted application parts.

17. CORRESPONDENCE ADDRESS

☒ Customer Number or Bar Code Label 000023703 or ☐ Correspondence address below
(Insert Customer No. or Attach bar code label here)

Name Dr. Robert Raymond
Boehringer Ingelheim Corporation
Address 900 Ridgebury Road
PO Box 368
City Ridgefield State CT Zip Code 06877
Country US Telephone 203/798-4864 Fax 203/798-4408

Name (Print/Type) Susan K. Pocchiari Registration No. (Attorney/Agent) 45,016
Signature Susan K. Pocchiari Date August 3, 2000

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any
comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office,
Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents,
Box Patent Application, Washington, DC 20231.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of: Konopitzky, Renate et al. Docket No.: 12/211
Serial No. : To be assigned) Art Unit:
Filed : August 3, 2000) Examiner:
For : Tumor-associated antigen (R11))

Jc784 U.S. PTO
09/631863
08/03/00

Assistant Commissioner for Patents
Washington, D.C. 20231

STATEMENT BY ATTORNEY UNDER 37 C.F.R. § 1.821(f)
REGARDING SEQUENCE LISTING

Sir:

Attorney for Applicants affirms that the information recorded in computer readable form is identical to the written sequence listing.

Date: August 3, 2000

Respectfully submitted,

Susan K. Pocchiari
Susan K. Pocchiari
Attorney for Applicant(s)
Reg. No. 45,016

Patent Department
Boehringer Ingelheim Corp.
900 Ridgebury Road
P.O. Box 368
Ridgefield, CT. 06877
Tel.: (203) 798-5648

TUMOR-ASSOCIATED ANTIGEN (R11)

Related Application

- 5 The benefit of prior provisional application Serial No. 60/154,161, filed September 15, 1999 is hereby claimed.

Field of the Invention

- 10 The invention relates to a novel tumor-associated antigen (TAA), derivatives and fragments thereof, antibodies thereto, and nucleic acids encoding the TAA and its derivatives and fragments. The invention further relates to the use of such molecules in the diagnosis and treatment or prevention of tumor diseases.

15 Background

- The immune system has the task of protecting the body from a number of different microorganisms and actively fighting these microorganisms. The importance of an intact immune system is apparent particularly in the case of inherited or acquired
- 20 immunodeficiencies. The use of prophylactic vaccine programmes proved in many cases to be an extremely effective and successful immunological intervention in the fight against viral or bacterial infectious diseases. It has also been found that the immune system is also involved to a large extent in eliminating tumour cells. Recognition of the tumour associated antigens (TAAs) by components of the immune system plays a crucial role. In
- 25 the broadest sense, any (peptidic or non-peptidic) component of a tumour cell which is recognised by an element of the immune system and leads to stimulation of an immune response can act as an immunogenic tumour antigen. Those tumour antigens which not only evoke an immunological reaction but also cause rejection of the tumour are of particular importance. The identification of specific antigens which are able to provoke an
- 30 immunological reaction of this kind constitutes a major step in developing a molecularly defined tumour vaccine. Although it is not yet clear which elements of the immune system are responsible for rejection of the tumour, there is nevertheless consensus that CD8-expressing cytotoxic T-lymphocytes (CTLs) play a major part (Coulie, 1997, Mol. Med. Today 3: 261-268). Particularly in those types of tumour (such as melanoma and kidney
- 35 carcinoma) which have a relatively high spontaneous remission rate, a correlation has been

- found between the clinical progress and the increased appearance of CD8⁺- and CD4⁺-T-cells (Schendel et al., 1993, J. Immunol. 151: 4209-4220; Mackensen et al., 1993, Cancer Res. 53: 3569-3573; Halliday et al., 1995, World J. Surg. 19: 352-358; Kawakami et al., 1995, J. Immunol. 154: 3961-3968; Kawakami et al., 1996, Med. 45: 100-108; Wang, 1997, Mol. Med. 3: 716-731; Celluzzi and Falo, 1998, J. Immunol. 160: 3081-3085). Specific CTL clones were obtained either from tumour-infiltrating lymphocytes (TIL) or peripheral mononuclear blood cells (PBMC) after co-cultivation with generally autologous tumour cells and cytokine stimulation *in vitro*. Both in animal models and in human cell culture systems cultivated *in vitro*, the T-cell response against tumour cells was increased by transfection of tumour cells with cytokines (van Elsas et al., 1997, J. Immunother. 20: 343-353; Gansbacher et al., 1990, J. Exp. Med. 172: 1217-1224; Tepper et al., 1989, Cell 57: 503-512; Fearon et al., 1990, Cell 60: 397-403; Dranoff et al., 1993, Proc. Natl. Acad. Sci. U. S. A 90: 3539-3543).
- In the light of the correlation between remission and the involvement of CD8⁺-T cells, the identification of tumour associated antigens (TAA) which are recognised by CD8-positive CTLs is a specific prime objective towards developing a tumour vaccine (Pardoll, 1998, Nature Medicine 4: 525-531; Robbins and Kawakami, 1996, Curr. Opin. Immunol. 5: 658-63). Whether other cell types of the immune system such as for example CD4⁺-T-helper cells play an important part is not yet clear; a number of studies with MAGE-3/HLA-A1 peptides in melanoma patients indicated this (Marchand et al., 1995, Int. J. Cancer 63: 883-885; Boon et al., 1998, Cancer Vaccine Week – International Symposium, New York, Oct 1998; abstract S01). In recent years a number of TAAs which are recognised by CTLs have been identified (Boon et al., 1994, Annu. Rev. Immunol. 12: 337-365; van den Eynde and van der Bruggen, 1997, Curr. Opin. Immunol. 9: 684-693).

T-cells recognise antigens as peptide fragments which are presented on the cell surfaces of MHC molecules (major histocompatibility complex, in man "HLA" = "human leukocyte antigen"). There are two types of MHC molecules: MHC-I molecules occur in most cells with a nucleus and present peptides (usually 8-10-mers) which are produced by proteolytic degradation of endogenous proteins (so-called antigen processing). Peptide: MHC-I complexes are recognised by CD8-positive CTLs. MHC-II molecules occur only on so-called "professional antigen-presenting cells" (APC) and present peptides of exogenous proteins which are absorbed and processed in the course of endocytosis by APC. Peptide: MHC-II complexes are recognised by CD4-helper-T cells. By interaction between the T-

cell receptor and peptide:MHC complex, various effector mechanisms may be triggered which lead to apoptosis of the target cell in the case of CTLs. This occurs if either the MHC (e.g. in the case of transplant rejection) or the peptide (e.g. in the case of intracellular pathogens) is recognised as foreign. In any case, not all the presented peptides meet the structural and functional requirements for effective interaction with T-cells (as described by Rammensee et al., 1995, Immunogenetics 41: 178-228 and hereinafter).

In principle, a number of methods of administration are possible for using TAAs in a tumour vaccine: the antigen can either be administered as a recombinant protein with suitable adjuvants or carrier systems or it may be given as cDNA coding for the antigen in plasmid (DNA vaccine; Tighe et al., 1998, Immunol. Today 19: 89-97) or viral vectors (Restifo, 1997). Another possibility is to use recombinant bacteria (e.g. listeria, salmonella) which recombinantly express the human antigen and have an adjuvant effect as a result of their additional components (Paterson, 1996, Curr. Opin. Immunol. 5: 664-669; Pardoll, 1998, Nature Medicine 4: 525-531). In all these cases, the antigen has to be processed and presented by so-called "professional antigen presenting cells" (APC). Another possibility is to use synthetic peptides (Melief et al., 1996, Curr. Opin. Immunol. 8: 651-657) which correspond to the equivalent T-cell epitopes of the antigen and are either loaded onto the APC from outside (Buschle et al., 1997, Proc. Natl. Acad. Sci. U.S.A. 94: 3256-3261; Schmidt et al., 1997, Proc. Natl. Acad. Sci. U.S.A. 94: 3262-3267) or absorbed by the APC and transferred intracellularly to the MHC I molecules. The most therapeutically efficient method of administration of a specified antigen is generally determined by clinical trials.

The antigens or epitopes thereof recognised by the tumour-specific CTLs include molecules which can come from any protein classes (e.g. transcription factors, receptors, enzymes; for a survey see Rammensee et al., 1995, Immunogenetics 41: 178-228; Robbins and Kawakami, 1996, Curr. Opin. Immunol. 8: 628-636). These proteins do not necessarily have to be located on the cell surface, as is necessary for recognition by antibodies. In order to act as a tumour specific antigen for recognition by CTLs or in order to be used for therapy, the proteins must meet certain conditions: first of all, the antigen should be expressed exclusively by tumour cells or should occur in so-called "critical" normal tissues not at all or only in smaller concentrations than in tumours. Critical normal tissues are essential tissues; an immune reaction directed against them would have severe, in some cases lethal consequences. Secondly, the antigen should be present not only in the

primary tumour but also in the metastases. Furthermore, with a view to broad clinical use of the antigen, it is desirable for it to be present in high concentrations in several types of tumour. One further precondition for the suitability of a TAA as an effective ingredient of a vaccine is the presence of T-cell epitopes in the amino acid sequence of the antigen;

5 peptides derived from the TAA should lead to an *in vitro/in vivo* T-cell response ("immunogenic" peptide). Another criterion for selecting a clinically broadly applicable immunogenic peptide is the frequency with which the antigen is encountered in a given population of patients.

10 The immunogenic tumour-associated antigens (TAAs), which have already largely been shown to have T-cell epitopes, can be divided into a number of categories, including viral proteins, mutated proteins, overexpressed proteins, fusion proteins formed by chromosomal translocation, differentiation antigens, oncofoetal antigens (Van den Eynde and Brichard, 1995, Curr. Opin. Immunol. 7: 674-681; van den Eynde and van der
15 Bruggen, 1997, Curr. Opin. Immunol. 9: 684-693).

The methods of identifying and characterising TAAs which form the starting point for the development of a tumour vaccine are based on the one hand on the use of CTLs which have already been induced in patients (cellular immune response) or antibodies (humoral
20 immune response), or are based on drawing up differential transcription profiles between tumours and normal tissues. In the former case, the immunological approach, patient CTLs are used for screening eukaryotic tumour-cDNA expression libraries which present the CTL-epitopes via MHC-I molecules (Boon et al., 1994, Annu. Rev. Immunol. 12: 337-365), whereas by using high affinity patient antisera prokaryotic cDNA expression
25 libraries, the presence of TAAs can be searched directly via immunoblot analysis of the individual plaques (Sahin et al., 1995, Proc. Natl. Acad. Sci. U.S.A. 92: 11810-11813). A combination of CTL reactivity and protein-chemical processes produces the isolation of peptides isolated from MHC-I from tumour cells, which are preselected by reactivity with patient CTLs. The peptides are washed out of the MHC-I complex and identified by mass
30 spectrometry (Falk et al., 1991, Nature 351: 290-296; Woelfel et al., 1994, Int. J. Cancer 57: 413-418; Cox et al., 1994, Science 264: 716-719). The approaches which use CTLs to characterise antigens involve substantial costs or are not always successful, owing to the need to cultivate and activate CTLs.

Methods of identifying TAAs which are based on comparing the transcription profile of normal and tumour tissue are many and varied; these include differential hybridisation, the establishing of subtraction cDNA banks ("representational difference analysis"; Hubank and Schatz, 1994, Nucleic Acids Res. 22: 5640-5648; Diatchenko et al., 1996, Proc. Natl. Acad. Sci. U.S.A. 93: 6025-6030) and the use of DNA chip technology or the SAGE method (Velculescu et al., 1995, Science 270: 484-487). In contrast to the above-mentioned immunological method using patient CTLs, when using molecular biological methods it is necessary to show that the potential antigen candidates discovered by this method are tumour-specific (tumour-associated) and do indeed have T-cell epitopes capable of triggering a cytotoxic T-cell response. In at least one case (NY-ESO/LAGE-1) an antigen was identified both by the use of patient sera and by RDA (Chen et al., 1997, Proc. Natl. Acad. Sci. U.S.A. 94: 1914-1918; Lethe et al., 1998, Int. J. Cancer 76: 903-908), and moreover CTL-epitopes of this antigen and a simultaneous spontaneous humoral and T-cell response were described in one patient (Jager et al., 1998, J. Exp. Med. 187: 265-270).

Summary of the Invention

The present invention relates to a new tumor-associated antigen designated R11. The invention further relates to R11 fragments and derivatives, nucleic acids encoding R11 and R11 fragments and derivatives, and antibodies and antibody fragments thereto which display one or more functional activities of the R11 protein such as specifically binding the R11 protein, inducing or augmenting an immune response (e.g., induction of CTLs, induction of antibodies), or treating or preventing cancer (e.g., reducing the volume or inhibiting the growth of a tumor that expresses R11).

Summary of the Figures

Fig. 1: Transcription of R11 in tumour tissues and normal tissues: Semi-quantitative RT-PCR

Fig. 2: Transcription of R11 in tumour tissues and normal tissues: Qualitative PCR

Fig. 3: Northern Blot analysis of R11 in normal tissues

Fig. 4: Transcription of R11: Qualitative RT-PCR from RNA from human tumour cell lines

Fig. 5: Modified region of the pCR3.1(+) vector.

- 5 The aim of the present invention was to provide a new tumour-associated antigen (TAA).

Detailed Description of the Invention

10 This objective was achieved by first establishing a cDNA subtraction library by RDA (representational difference analysis) between a cell line derived from a pancreas carcinoma patient and a pool of 11 different normal tissues. In order to generate the cDNA fragments of "tester" and "driver" required for the subtractive hybridisation, in a departure from the original method (Diatchenko et al., 1996, Proc. Natl. Acad. Sci. U. S. A. 93: 6025-6030) a mixture of 6 different restriction enzymes was used. The use of a mixture of
15 different restriction enzymes which require 6 base pairs as the recognition sequence has the following advantages over the original method (Diatchenko et al., 1996, Proc. Natl. Acad. Sci. U. S. A. 93: 6025-6030): a) by selecting two restriction enzymes the recognition sequences of which are represented by combinations of 6 of the bases A/T (e.g. Ssp I: AATATT) or C/G (e.g. Nae I: GCCGGC) or A/C/G/T (e.g. EcoR V: GATATC), both GC-
20 and AT-rich regions of a gene are cut in the same way, thus permitting homogeneous representation of the entire gene region as restriction fragments; b) in addition, this makes it possible to obtain larger cDNA fragments of the candidate gene, on a statistical average (about 800 bp), which is in turn highly advantageous in the subsequent analysis (sequencing and annotation) and cloning of the "full-size" cDNA. In the original method
25 (Diatchenko et al., 1996, Proc. Natl. Acad. Sci. U. S. A. 93: 6025-6030) a restriction enzyme (Rsa I) recognising only 4 bases was used, which leads to an average fragment length of 256 bp and cannot specifically process CG- or AT-rich regions. In order to do justice to the hybridisation kinetics changed by the longer insert cDNA fragments, the PCR procedure was modified as described in Example 2.

30 In order to select the antigens which were overexpressed in the tumour, the cDNA clones obtained were first separated and a basic glycerol culture, a plasmid preparation and an insert-representing collection of the PCR fragments were established therefrom in a 96-well dish format. First, 50 randomly selected cDNA fragments of the 3450 clones of the
35 subtractive cDNA library of the pancreatic carcinoma were sequenced in order to select the

antigens which are overexpressed in the tumour and compared with sequences available in data banks. Among the genes annotated there were 12 unknown ones, for which there were EST entries (expressed sequence tags) in the data bank. One clone, R11, by its preferential presence in foetal tissue, indicated an EST profile which was suitable for possible use as a TAA. Further investigations using semi-quantitative RT-PCR and Northern Blot analysis confirmed the preferred expression in various tumour (carcinoma of the breast, kidney cells and pancreas) and immunoprivileged tissues (testis, placenta and adrenal glands) and little or no expression in normal tissues. Moreover, it can be concluded from the data obtained by Northern Blot experiments that the R11 transcript is about 7.5 kb long and that splice variants or homologous genes may possibly exist in the adrenal glands.

The human R11-cDNA was cloned from testis; the sequence obtained is shown in SEQ ID NO:1. The sequence of R11 shows no identity or homology with any known gene at either the nucleotide or protein level. The R11-cDNA obtained within the scope of the present invention contains two separate open reading frames for a protein 401 amino acids long (SEQ ID NO:2) and for a protein 357 amino acids long (SEQ ID NO:3). The R11-cDNA cloned within the scope of the present invention has a length of 6582 bp, whilst the presence of a PolyA tail at the 3'-end of the sequence is an indication of the completeness of the cDNA in this region.

On the basis of the data obtained within the scope of the present invention, it cannot be ruled out that 5' from the sequenced cDNA there is another ATG which constitutes the start ATG for the first open reading frame (R11-ORF-1); in this case, the cDNA present contains the region coding for the C-terminal section of R11-ORF-1 at the 5' end. Information as to the 5' end and a possible coding DNA sequence section located further upstream can be obtained by standard methods of molecular biology, e.g. by 5'-RACE (rapid amplification of cDNA ends). In this method, RNA, preferably mRNA, is reverse transcribed from cells or tissues in which R11 is transcribed (e.g. tissue from carcinoma of the breast, kidney cells or pancreas) and then ligated with an adapter of known sequence. A PCR with an adapter primer (binding specifically to the adapter at the 5'-end of the cDNA) and an R11-specific primer (e.g. SEQ ID NO: 26) allows corresponding R11 fragments to be amplified. These PCR products can be cloned by standard methods and characterised, particularly by DNA sequencing, as described in Example 6.

An alternative method of characterising the 5'-end is by screening cDNA libraries by hybridisation with DNA probes which are specific for R11 or analysis of cDNA expression libraries with antisera.

- 5 If the screening of cDNA libraries does not achieve the desired outcome, on account of limitations of procedure, e.g. inefficient reverse transcription caused by marked secondary structures of the RNA, genomic libraries can be searched by, for example, isolating clones, as in the screening of cDNA libraries, by hybridising with DNA probes specific for R11, said clones containing the sequence information located upstream of the 5'-end of the
10 cDNA obtained, e.g. the promoter region of R11.

In the course of total cloning of the R11-cDNA it is possible to establish whether the open reading frame of R11-ORF-1 obtained in the region of the cDNA fragment present has a continuation in the 5' region and/or whether there are alternative reading frames.

- 15 The cDNA isolated within the scope of the present invention has the nucleotide sequence given in SEQ ID NO:1; it is to be assumed (see above) that it codes for the C-terminal portion of a tumour-associated antigen (TAA) designated R11-ORF-1 and for another protein which is represented by the second reading frame (R11-ORF-2).

- 20 The two proteins of the two reading frames expressed by the isolated cDNA have the amino acid sequence shown in SEQ ID NOs:2 and 3, respectively.

- 25 In a first aspect, the present invention relates to an isolated DNA molecule which has the nucleotide sequence shown in SEQ ID NO:1 or a polynucleotide which hybridises with this DNA molecule under stringent conditions.

- By 'stringent conditions' is meant, for example: incubation overnight at 65°C – 68°C with 6xSSC (1x SSC = 150 mM NaCl, 15mM trisodium citrate), 5xDenhardt's solution, 0.2%
30 SDS, 50µg/ml salmon sperm DNA, followed by washing twice for 30 min with 2xSSC, 0.1% SDS at 65°C, once for 30 min with 0.2xSSC, 0.1% SDS at 65°C and optionally finally rinsing with 0.1xSSC, 0.1%SDS at 65°C.

In another aspect the present invention relates to an isolated DNA molecule which contains a polynucleotide of the sequence shown in SEQ ID NO: 1 as a partial sequence or which

contains a polynucleotide which hybridises with a polynucleotide of this sequence under stringent conditions.

The nucleic acids or fragments thereof according to the invention code for polypeptides designated R11-ORF-1 and R11-ORF-2, whilst R11-ORF-2 has the amino acid sequence shown in SEQ ID NO:3 and R11-ORF-1 has the amino acid sequence shown in SEQ ID NO:2 or contains it; or for protein fragments or peptides derived from R11-ORF-1 or R11-ORF-2. This, therefore, includes DNA molecules which comprise deviations from the sequence shown in SEQ ID NO:1 as a result of the degeneration of the genetic code.

In another aspect the present invention relates to the tumour-associated antigens designated R11-ORF-1 and R11-ORF-2, whilst in the case of R11-ORF-1, if there is an extension of the existing open reading frame in the 5' direction, the amino acid sequence given in SEQ ID NO:2 for R11-ORF-1 is a partial sequence. The proteins with the sequences shown in SEQ ID NO: 2 and 3 are products which are translated by a transcript approximately 7.5 kb in size, or which are translated by transcripts about 3.8 kb and 2.3 kb in size which are derived from splice variants of the 7.5 kb transcript such as may be found in the adrenal tissues, or from transcripts of the genes homologous thereto.

The amino acid sequences shown in SEQ ID NO:2 and 3 may have deviations, e.g. those caused by the replacement of amino acids, insofar as the R11 derivatives ("R11", unless otherwise specified, herein denotes R11-ORF-1 and/or R11-ORF-2) have the immunogenic properties desirable for use in a tumour vaccine.

The natural amino acid sequence of R11-ORF-1 or R11-ORF-2 can optionally be modified by replacing individual amino acids in an R11 CTL-epitope in order to achieve an increase in the affinity of R11 peptides to MHC-I molecules compared with the natural R11 CTL-epitope, and thus bring about increased immunogenicity and finally greater reactivity to tumours. Modifications in the region of the R11 epitopes may be carried out on the whole R11 protein (this is processed by the APCs to form the corresponding peptides) or on larger R11 protein fragments or on R11 peptides (cf. below).

According to another aspect, the present invention relates to immunogenic polypeptide fragments and peptides derived from R11-ORF-1 or R11-ORF-2. The latter are hereinafter

referred to as R11 peptides. A first group are the R11 peptides which trigger a humoral immune response (induction of antibodies). Such peptides are selected portions of R11-ORF-1 or R11-ORF-2 (at least 6 amino acids) which can be determined by so-called prediction algorithms such as for example the surface probability plot (Emini et al., 1985 J. Virol. 55: 836-839), the hydrophobicity plot (Kyte and Doolittle, 1982 J. Mol. Biol. 157: 105-132) and the antigenic index (Jameson and Wolf, 1988 Comput. Appl. Biosci. 4: 181-186).

Also included are all those peptides which are derived from the N-terminal region of R11 which is optionally obtained in the course of further cloning.

It is known that tumour-associated antigens may have tumour-specific mutations which contribute to an immunological distinction between tumour and normal tissue (Mandrizzato et al., 1997, J Exp Med 186: 785-793; Hogan et al., 1998, Cancer Res 58: 5144-5150; Gaudi et al., 1999, J Immunol 162: 1730-1738; Wolfel et al., 1995, Science 269: 1281-1284). In order to detect the presence of tumour-specific R11-ORF-1 or R11-ORF-2 mutations, the R11 cDNA is cloned from one or more different tumours, appropriately using probes of the cDNA isolated from testis according to the invention, and the sequences obtained are compared with normal tissue R11-cDNA. It is to be expected that tumour R11 peptides from a fragment of sequence which is mutated compared with normal tissue R11 will exhibit greater immunogenicity compared with normal tissue R11 peptides from the corresponding fragment. Thus, according to another aspect, the present invention relates to R11 peptides derived from regions of a tumour-expressed R11-ORF-1 or R11-ORF-2 which contain tumour-specific mutations.

For therapeutic use, R11 peptides are administered directly or in modified form (e.g. coupled to KLH (keyhole limpet hemocyanine)) and the production of antibodies is determined by common immunological assays, e.g. ELISA.

Other R11 peptides which are preferred within the scope of the present invention are those which are presented by MHC-molecules and produce a cellular immune response. There are two types of MHC-molecules, namely MHC-I molecules which are recognised by CD8-positive CTLs and MHC-II molecules which are recognised by CD4-positive T-helper cells.

In order for a peptide to trigger a cellular immune response, it must bind to an MHC-molecule, whilst the patient to be treated must have the MHC molecule in his or her repertoire. Determining the MHC-subtype of the patient thus constitutes one of the essential prerequisites for effective use of a peptide in this patient, with a view to triggering
 5 a cellular immune response.

The sequence of an R11 peptide to be used therapeutically is determined by the MHC-molecule in question in terms of the anchor amino acids and length. Defined anchor positions and length guarantee that a peptide fits the peptide binding groove of the MHC-
 10 molecule of the patient in question. The result of this is that the immune system is stimulated and a cellular immune reaction is produced which is directed against the tumour cells of the patient, if a peptide derived from a tumour antigen is used.

Immunogenic R11 peptides may be identified by known methods; one of the basic
 15 conditions is the correlation between MHC-binding and CTL-induction.

Thus, since the sequence of immunogenic peptides can be predicted on the basis of its peptide binding motif, R11 peptides which constitute CTL-epitopes can be identified and synthesised on the basis of the R11 protein sequence. Various methods are available for
 20 doing this, which are used to identify CTL-epitopes of known protein antigens; e.g. the method described by Stauss et al., 1992, Proc. Natl. Acad. Sci. U.S.A 89: 7871-7875 for identifying T-cell epitopes in human papilloma virus.

The allele-specific requirements of each MHC-I allele product with regard to a peptide
 25 which binds to the MHC-molecule and is presented thereby have been assembled as a motif (e.g. Falk et al., 1991, Nature 351: 290-296). Up until now, a large number of both MHC-peptide motifs and MHC-ligands have become known. A suitable method, within the scope of the present invention, for searching for epitopes of a known protein which fits a specific MHC-I molecule is described in a survey by Rammensee et al., 1995,
 30 Immunogenetics 41: 178-228. It comprises the following steps: first, the protein sequence is searched for fragments which correspond to the anchor motif, whilst certain variations are possible with regard to peptide length and anchor occupation. If, for example, a motif prescribes a 9-mer with Ile or Leu at the end, 10-mers with a corresponding C-terminus can also be considered, as can peptides with other aliphatic groups such as Val or Met at the C-
 35 terminus. In this way a number of peptide candidates is obtained. These are searched for

the presence of as many anchor groups as possible which they have in common with known ligands and/or to see whether they have groups which are "preferred" for various MHC-molecules (according to the Table by Rammensee et al., 1995, Immunogenetics 41: 178-228). In order to exclude weakly binding peptides, binding assays are preferably

5 carried out. If the requirements for the peptide binding for specific MHC-molecules are known, the peptide candidates can also be searched for non-anchor groups which have a negative or positive effect on the binding or which indeed make it possible at all (Ruppert et al., 1993, Cell 74: 929-937). However, with this method, it should be borne in mind that the peptide binding motif is not the sole deciding factor when searching for natural ligands;

10 other aspects, e.g. enzyme specificity during antigen processing, also contribute to the identity of the ligand, in addition to the specificity of the MHC-binding. One method which takes account of these aspects and which is suitable for identifying immunogenic R11 peptides within the scope of the present invention was used *inter alia* by Kawakami et al., 1995, J. Immunol. 154: 3961-3968 for identifying gp100 epitopes on the basis of

15 known HLA-A*0201 motifs.

The peptides may also be selected for their ability to bind to MHC-II molecules. The MHC-II binding motif which extends over nine amino acids has a higher degree of degeneration in the anchor positions than the MHC-I binding motif. Methods have

20 recently been developed, based on X-ray structural analysis of MHC-II molecules, which allow accurate analysis of the MHC-II binding motifs and, based on that, variations in the peptide sequence (Rammensee et al., 1995, Immunogenetics 41:178-228, and the original literature cited therein). Peptides which bind to MHC-II molecules are typically presented to the CD4-T cells by dendritic cells, macrophages or B-cells. The CD4-T-cells in turn

25 then activate CTLs directly in sequence by the release of cytokine, for example, and increase the efficiency of antigen presentation by APC (dendritic cells, macrophages and B-cells).

Recently, databanks and prediction algorithms have become available which allow more

30 reliable prediction of peptide epitopes which bind to a specific MHC molecule.

Within the scope of the present invention, using the algorithm described by Parker et al., 1994, J. Immunol. 152: 163 and Rammensee et al., 1995, Immunogenetics 41:178-228, candidate peptides of the C-terminal fragment of R11 have been identified for the most

35 important HLA-types, especially for HLA-A1, -A*0201, -A3, -B7, -B14 and -B*4403,

which can be expected to bind to the corresponding HLA molecules and thus constitute immunogenic CTL-epitopes; the peptides discovered are listed in Table 1. Similarly, possibly using other algorithms which take account of the different characteristics of the peptides (hydrophobicity, charge, size) or requirements made of the peptides, such as the
 5 3D structure of the HLA-molecule, it is possible to find other potential peptide epitopes; this also applies to peptide epitopes of other HLA types.

After selecting R11-peptide candidates using the methods described, their MHC-binding is tested by peptide binding assays. First, the immunogenicity of the peptides with good
 10 binding properties is determined (stability of the peptide-MHC interaction correlates in most cases with immunogenicity; van der Burg et al., 1996, J. Immunol. 156: 3308-3314). In order to determine the immunogenicity of the selected peptide or peptide equivalent, methods may be used as described, for example, by Sette et al., 1994, J. Immunol. 153: 5586-5592 combined with quantitative MHC-binding assays. Alternatively, the
 15 immunogenicity of the selected peptide may be tested by *in vitro* CTL-induction using known methods (as described hereinafter for *ex vivo* CTL-induction). The principle of the method, carried out in several steps, for selecting peptides which are capable of triggering a cellular immune response is described in WO 97/30721, the contents of which are hereby expressly referred to. A general strategy for obtaining efficient immunogenic peptides
 20 which is suitable within the scope of the present invention has also been described by Schweighoffer, 1997, Onc. Res. 3: 164-176.

Instead of using the original peptides which fit the binding groove of MHC-I or MHC-II molecules, i.e., peptides which are derived unaltered from R11, variations may be carried
 25 out, adhering to the minimum requirements regarding anchor positions and length specified on the basis of the original peptide sequence, provided that these variations not only do not impair the effective immunogenicity of the peptide which is made up of its binding affinity to the MHC-molecule and its ability to stimulate T-cell receptors, but preferably enhance it. In this case, artificial peptides or peptide equivalents are thus used which are designed
 30 to correspond to the requirements regarding binding ability to an MHC-molecule.

Peptides modified in this way are referred to as "heteroclitic peptides". They may be obtained by the following methods:

First of all, the epitopes of MHC-I or MHC-II ligands or variations thereof are undertaken, e.g. using the principle described by Rammensee et al., 1995, Immunogenetics 41:178-228. The length of the peptide preferably corresponds to a minimum sequence of 8 to 10 amino acids with the necessary anchor amino acids, if the peptide is being matched to MHC-I molecules.

If desired, the peptide may also be extended at the C- and/or N-terminus provided that this extension does not affect the ability to bind to the MHC-molecule and the extended peptide can be cellularly processed down to the minimum sequence.

The modified peptides are then investigated for their recognition by TILs (tumour infiltrating lymphocytes), for CTL-induction and for increased MHC-binding and immunogenicity, as described by Parkhurst et al., 1996, J. Immunol 157: 2539-2548 and Becker et al., 1997, J. Immunol. Methods 203: 171-180.

Another method of finding peptides with greater immunogenicity than that of the natural R11 peptides, which is suitable for the purposes of the present invention, consists in screening peptide libraries with CTLs which recognise the R11 peptides naturally occurring on tumours, as described by Blake et al., 1996, J. Exp. Med. 184: 121-130; in connection with this it is proposed to use combinatorial peptide libraries in order to design molecules which imitate tumour epitopes recognised by MHC-I-restricted CTLs.

The R11 polypeptides according to the present invention or immunogenic fragments or peptides derived therefrom may be produced recombinantly or by peptide synthesis, as described in WO 96/10413, the disclosure of which is hereby referred to. For recombinant production, the corresponding DNA molecule is inserted by standard methods in an expression vector, transfected into a suitable host cell, the host is cultivated under suitable expression conditions and the protein is purified. Conventional methods may be used for the chemical synthesis of R11 peptides, e.g. automatic peptide synthesisers which are commercially available.

Alternatively to natural R11 peptides or heteroclitic peptides, it is also possible to use substances which imitate such peptides, e.g. "peptidomimetics" or "retro-inverse peptides". In order to test these molecules with regard to their therapeutic use in a tumour vaccine the

same methods are used as described above for the natural R11 peptides or R11 peptide equivalents.

5 The two TAAs designated R11-ORF-1 and R11-ORF-2 according to the present invention and the protein fragments, peptides or peptide equivalents or peptidomimetics derived therefrom may be used in cancer therapy, e.g. in order to induce an immune response to tumour cells which express the corresponding antigen determinants. They are preferably used for the treatment of R11-ORF-1- and/or R11-ORF-2-positive tumours, particularly in carcinoma of the breast, kidney cells and pancreas.

10

The immune response in the form of induction of CTLs can be achieved *in vivo* or *ex vivo*.

15 In order to induce CTLs *in vivo*, a pharmaceutical composition containing as active component the TAAs R11-ORF-1 and/or R11-ORF-2 or fragments or a peptide or peptides derived therefrom, is administered to a patient suffering from a tumoral disease associated with the TAA, whilst the quantity of TAA (peptide) must be sufficient to obtain an effective CTL response to the antigen-bearing tumour.

20 Thus, according to another aspect, the invention relates to a pharmaceutical composition for parenteral, topical, oral or local administration. Preferably, the composition is used parenterally, e.g. for subcutaneous, intradermal or intramuscular application, containing as active component the TAAs R11-ORF-1 and/or R11-ORF-2 or fragments or peptide(s) derived therefrom. The R11-TAAs/peptides are dissolved or suspended in a pharmaceutically acceptable, preferably aqueous, carrier. The composition may also
25 contain conventional adjuvants such as buffers etc. The R11-TAAs/peptides may be used on their own or in conjunction with adjuvants, e.g. incomplete Freund's adjuvant, saponines, aluminium salts or, in a preferred embodiment, polycations such as polyarginine or polylysine. The peptides may also be bound to components which aid CTL induction or CTL activation, e.g. T-helper peptides, lipids or liposomes, or they are administered
30 together with these substances and/or together with immunostimulant substances, e.g. cytokines (IL-2, IFN- γ). Methods and formulations which are suitable for the preparation and administration of the pharmaceutical composition according to the invention are described in WO 95/04542 and WO 97/30721, the disclosures of which are hereby referred to.

35

R11 polypeptide fragments or R11 peptides may also be used to trigger a CTL response *ex vivo*. An *ex vivo* CTL response to a tumour which expresses the two possible proteins of R11 is induced by incubating the CTL-precursor cells together with APCs and R11 peptides or R11 protein. The activated CTLs are then allowed to expand, whereupon they are re-administered to the patient. Alternatively, APCs may be loaded with R11 peptides, which may lead to efficient activation of cellular immune reactions against R11 positive tumours (Mayordomo et al., 1995, Nature Medicine 1: 1297-1302; Zitvogel et al., 1996, J. Exp. Med. 183: 87-97). One suitable method of loading peptides onto cells, e.g. dendritic cells, is disclosed in WO 97/19169.

In one embodiment of the invention a combination of several different R11 peptides or R11 peptide equivalents is used. In another embodiment, R11 peptides are combined with peptides derived from other TAAs. The choice of peptides for such combinations is made in the light of detecting different MHC-types in order to cover the broadest possible patient population, and/or it is aimed at the broadest possible spectrum of indications, by combining peptides from several different tumour antigens. The number of peptides in a pharmaceutical composition can fluctuate over a wide range, but typically a clinically usable vaccine contains 1 to 15, preferably 3 to 10 different peptides.

The peptides according to the invention may also be used as diagnostic reagents. For example, the peptides may be used to test the response of a patient to the humoral or cellular immune response evoked by the immunogenic peptide. This provides a possibility of improving a treatment procedure. For example, depending on the form of administration (peptide, total protein or DNA vaccine) of the TAA, the increase of precursor T-cells in the PBLs which show reactivity against the defined peptide epitope can be investigated (Robbins and Kawakami, 1996, Curr. Opin. Immunol. 8: 628-636 and the references cited therein). Moreover, the peptides or the total protein or antibodies directed against the TAA may be used to characterise the progression of an R11-ORF-1- or R11-ORF-2-positive tumour (e.g. by immunohistochemical analyses of primary tumour and metastases). A strategy of this kind has already proved successful in many cases, e.g. detecting the oestrogen receptor as the basis for deciding on endocrine therapy in breast cancer; c-erbB-2 as the relevant marker in the prognosis and course of therapy in breast cancer (Ravaioli et al., 1998, Cell. Prolif. 31: 113-126; Revillion et al., 1998, Eur. J. Cancer 34: 791-808); PSMA (prostate specific membrane antigen) as a marker for epithelial cells of prostate carcinoma in the serum or by using a ¹¹¹In-labelled monoclonal

antibody against PSMA in immunoscintigraphy on prostate carcinoma (Murphy et al., 1998, Cancer 83: 2259-2269 and references included therein); CEA (carcinoembryonic antigen) as a serological marker for the prognosis and progression in patients suffering from colorectal carcinoma (Jessup and Loda, 1998, Semin. Surg. Oncol. 15: 131-140).

5

Of the DNA molecules according to the invention defined above, those which lead, by mutation, to an exchange of amino acids in the protein sequence shown in SEQ ID NO:2 or 3, if they code for an R11 derivative or fragments or peptides with the immunogenic properties which are desirable for their use as tumour vaccines, are also included.

10

The R11 DNA molecules of the present invention or the corresponding RNAs which are also a subject of the present invention are used, like the (poly)peptides coded by them, for immunotherapy of cancer diseases.

15

In one embodiment of the invention, DNA molecules are used which code for natural R11 polypeptides. Alternatively to the natural R11 cDNA or fragments thereof it is possible to use modified derivatives. These comprise sequences with modifications which code for a protein (fragment) or peptides with greater immunogenicity, whilst the same considerations apply to modifications at the DNA level as apply to the peptides described above. Another type of modification is the lining up of numerous sequences coding for immunologically relevant peptides like a string of beads (Toes et al., 1997, Proc. Natl. Acad. Sci. U.S.A. 94: 14660-14665). The sequences may also be modified by the addition of auxiliary elements, e.g. functions, which ensure more efficient release and processing of the immunogen (Wu et al., 1995, Proc. Natl. Acad. Sci. U.S.A. 92: 11671-11675). For example, the processing and hence the presentation and finally the immunogenicity of the antigen can be increased by the addition of a locating sequence in the endoplasmatic reticulum ("ER targeting sequence").

20

25

In another aspect, the present invention relates to a recombinant DNA molecule which contains the R11-DNA according to SEQ ID NO:1 or a partial sequence, particularly the sequence coding for the polypeptide R11-ORF-1 or R11-ORF-2.

30

The R11 DNA molecules of the present invention may be administered, preferably in recombinant form as plasmids, directly or as part of a recombinant virus or bacterium. In

theory, any method of gene therapy may be used for immunotherapy of cancer based on DNA ("DNA vaccine") on R11-DNA, both *in vivo* and *ex vivo*.

5 Examples of *in vivo* administration are the direct injection of "naked" DNA, either by intramuscular route or using a gene gun, which has been shown to lead to the formation of CTLs against tumour antigens. Examples of recombinant organisms are vaccinia virus, adenovirus or listeria monocytogenes (a summary was provided by Coulie, 1997, Mol. Med. Today 3: 261-268). Moreover, synthetic carriers for nucleic acids such as cationic lipids, microspheres, micropellets or liposomes may be used for *in vivo* administration of
10 nucleic acid molecules coding for R11 peptide. As with peptides, different adjuvants which enhance the immune response may also be administered, e.g. cytokines, either in the form of proteins or plasmids coding for them. The application may optionally be combined with physical methods, e.g. electroporation.

15 An example of *ex vivo* administration is the transfection of dendritic cells as described by Tuting, 1997, Eur. J. Immunol. 27: 2702-2707, or other APCs which are used as cellular cancer vaccines.

20 Thus, according to another aspect, the present invention relates to the use of cells which express R11, either *per se* or, in optionally modified form, after transfection with the corresponding coding sequence, in order to produce a cancer vaccine.

In another aspect, the invention relates to antibodies against R11-ORF-1 or R11-ORF-2 (hereinafter 'anti-R11-antibodies') or fragments thereof. Polyclonal anti-R11-antibodies
25 are conventionally obtained by immunising animals, particularly rabbits, by injecting the antigen or fragments thereof and subsequently purifying the immunoglobulin.

Monoclonal anti-R11-antibodies may be obtained by standard procedures following the principle described by Köhler and Milstein, 1975 Nature 265: 495-497, by immunising
30 animals, particularly mice, then immortalising antibody-producing cells from the immunised animals, e.g. by fusion with myeloma cells, and screening the supernatant of the hybridomas obtained by immunological standard assays for monoclonal anti-R11-antibodies. For therapeutic or diagnostic use in humans, these animal antibodies may optionally be chimerised in the conventional way (Neuberger et al., 1984 Nature 312: 604-

608, Boulianne et al., 1984 Nature 312: 643-646) or humanised (Riechmann et al., 1988, Nature 332: 323-327, Graziano et al., 1995, J. Immunol. 155: 4996-5002).

Human monoclonal anti-R11-antibodies (or fragments thereof) may also be obtained from
 5 so-called phage display libraries (Winter et al., 1994, Annu. Rev. Immunol. 12: 433-455, Griffiths et al., 1994, EMBO J. 13: 3245-3260, Kruif et al., 1995, Proc. Natl. Acad. Sci. USA 92: 3938-3942, McGuinness et al., 1996, Nature Biotechnol. 14: 1149) and by means of transgenic animals (Brüggemann et al., 1996 Immunol. Today 17: 391-397, Jakobovits et al., 1995, Curr. Opin. Biotechnol. 6: 561-566).

10 The anti-R11-antibodies according to the invention may be used in immunohistochemical analyses for diagnostic purposes.

In another aspect, the invention relates to the use of R11-ORF-1- and R11-ORF-2-specific
 15 antibodies for selectively bringing any desired substances to or into a tumour which expresses R11-ORF-1 and/or R11-ORF-2. Examples of such substances are cytotoxic agents or radioactive nuclides the activity of which consists in damaging the tumour *in situ*. Because of the tumour-specific expression of R11-ORF-1 or R11-ORF-2, no or very few side effects can be expected. According to another aspect, substances for detecting
 20 tumours which express R11 may be used, with the aid of R11-ORF-1 and/or R11-ORF-2 antibodies. This is useful for the diagnosis and evaluation of the treatment. Therapeutic and diagnostic uses of, are described in WO 95/33771.

The TAAs designated R11-ORF-1 and R11-ORF-2 according to the present invention and
 25 the protein fragments, peptides or peptide equivalents or peptidomimetics derived therefrom may be used in cancer therapy, e.g. to induce an immune response to tumour cells which express the corresponding antigen determinants. They are preferably used for the treatment of R11-ORF-1- and/or R11-ORF-2-positive tumours, particularly in carcinoma of the breast, kidney cells and pancreas.

30 In another application, R11 may be used as the target molecule of targeted chemotherapy.

By chemotherapy is meant the therapeutic administration of substances which have either a
 35 cytostatic or cytotoxic-cytolytic activity by interfering with the metabolism of malignant cells, their signal transduction and their cell division processes.

In principle, these chemotherapeutic agents develop their activity in all dividing cells; tumour cells, however, show greater sensitivity to these substances than healthy cells, as it is mainly strongly proliferating cells which are affected.

5

The prerequisite for the use of the tumour-associated R11 as a target for the chemotherapy is - unlike the immunological therapeutic approaches mentioned above - knowledge of the function of the R11 proteins R11-ORF-1 and R11-ORF-2 or the gene coding therefor.

- 10 The first step in the so-called 'downstream' functional analysis of R11 is conveniently a bioinformatic analysis which points the way for the experimental validation of R11 as a target for the chemotherapy.

The bioinformatic concepts based on similarity and modular structure constitute an
 15 essential basis for this analysis. Established bioinformatic aids for determining similarities are BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>, Altschul et al., 1997, Nucleic Acids Res. 25: 3389-3402) or FASTA (Pearson & Lipman, 1988, Proc. Natl. Acad. Sci. USA 85: 2444-2448), the specialised data banks such as Pfam (<http://www.sanger.ac.uk/Pfam>, Bateman et al., 2000, Nucleic Acids Res. 28: 263-266) and SMART ([http://smart.embl-](http://smart.embl-heidelberg.de)
 20 [heidelberg.de](http://smart.embl-heidelberg.de), Schultz et al., 2000, Nucleic Acids Res. 28: 231-234), which take account of domain structures. To refine the analysis, applications such as: Clustal (<http://www2.ebi.ac.uk/clustalw>, Higgins et al., 1996, Methods Enzymol. 266: 383-402); *HMMer* (<http://hmmer.wustl.edu>, Durbin et al, 1998, Cambridge University Press; PSI-BLAST (Altschul et al., 1997, Nucleic Acids Res. 25: 3389-3402) and the PROSITE data
 25 bank (<http://www.expasy.ch/prosite>, Hofmann et al., 1999, Nucleic Acids Res. 27: 215-219) may be used. Statistical methods of analysis which are not based on homologies make it possible to predict other structure- and function-related properties such as the secondary structure and the occurrence of transmembrane segments and helix-turn-helix motifs. Methods of predicting the secondary structure of proteins are available;
 30 particularly worth mentioning is *Jpred* (<http://barton.ebi.ac.uk/servers/jpred.html>, Cuff et al., 1998, Bioinformatics 14: 892-893). The prediction of the secondary structure may form the basis for functional hypotheses, e.g. if the structure of the presumed homologue is known.

Subsequently, R11 is subjected to a biochemical and biological analysis.

After the sequence analysis described above has been carried out, R11-ORF-1 and R11-ORF-2 are subjected to biochemical and biological analysis. The choice of the methods used for the further analyses depends on the result of the bioinformatic analysis carried out.

One example of functional analysis is the analysis of partly theoretically derived proteins of chromosome III of the yeast genome. (In such an analysis it was possible to predict more than 70% of the gene functions by the use of the bioinformatics, some of which were confirmed experimentally (Bork et al., 1992, *Nature Genetics* 18: 313-318; Sharp et al., 1993, *Nucleic Acids Res.* 21: 179-183 and Koonin et al., 1994, *EMBO J.* 13: 493-503).

In all the studies to be carried out it is important to preselect those domains of the protein of unknown function which is to be analysed which have a striking structural complexity, as limited structural information (e.g. globular regions) does not contribute to any major information content. An extensive summary of examples of successful predictions of function on the basis of protein sequences has been published in *Nature Genetics* by Bork and Koonin (Bork and Koonin, 1998, *Nature Genetics* 18: 313-318).

In the functional analysis of R11-ORF-1 carried out within the scope of the present invention it was established that, according to bioinformatic analysis, it is a protein which belongs to the family of the zinc finger-containing transcription factors.

By means of suitable experiments such as, for example, mobility shift, South-Western, UV-crosslinking, etc., it is possible to demonstrate a direct and/or indirect interaction with nucleic acids, particularly in promoter regions. Suitable methods for this are known from the literature (e.g. Ausubel et al., 1994, Vol. 1 and 2 "Current Protocols in Molecular Biology", John Wiley & Sons., Inc.).

For the first 280 amino acids of the protein derived from the R11-ORF-2 region it was possible to show a clear homology with a retroviral pol polyprotein. Thus, R11-ORF-2 might be a possible retrotransposon. Once the function of R11-ORF-1 or R11-ORF-2 is established, the significance of the R11 gene and its function or the function of the proteins coded thereby for the occurrence of tumours is analysed. This may be demonstrated, for example, by proliferation assays *in vitro* or in animal models using tumour cells which overexpress the gene under investigation (constitutively or inducibly) and as a control express it either in deleted (inactive) form or down-regulate it by antisense (cf. e.g. Grosveld and Kollias, 1992, Transgenic Animals, Academic Press).

10

R11 can be used in screening assays for identifying substances which modulate, especially inhibit, the activity of R11-ORF-1 or R11-ORF-2. In one embodiment an assay of this kind might consist, for example, of introducing R11-ORF-1 or R11-ORF-2 or an active fragment thereof into cells which react to the activity of R11 with proliferation or expressing the corresponding R11-cDNA-fragment in the cell and determining the proliferation of the cells in the presence and absence of a test substance.

15

One example of test cells is cells with a low division rate, e.g. primary cells which have no endogenous R11. In order to establish the suitability of cells for a screening assay, they are transformed with R11-cDNA, cultured and tested with standard assays, e.g. thymidine incorporation, for their ability to proliferate. On the basis of a significant increase in their ability to proliferate after R11 expression, they may be used as test cells, e.g. in High Throughput Screening Proliferation Assays. Examples of proliferation assays in the High Throughput format, e.g. based on the MTS assay, are described in WO 98/00713.

20

R11 inhibitors with a proliferation-inhibiting activity can be used to treat tumours with powerful R11 expression, particularly in carcinoma of the breast, kidney cell or pancreas.

In another aspect, the invention relates to a kit comprising in one or more containers a molecule consisting of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:3 or a protein fragment derived from R11 or a derivative thereof; and in

25

another container, an antibody that specifically binds to a TAA designated R11 or a fragment or derivative thereof.

Example 1

5

Cell culture of the cell line MZ.PC2 m7#1 B7.1#3 derived from a human pancreatic carcinoma and isolation of the poly A⁺ RNA

10 The cell line MZ.PC2 m7#1 B7.1#3 is derived from a human pancreatic carcinoma (MZ.PC2); it was obtained as follows: First, the tumour cells were passaged once through the mouse and a clone was selected for further study (MZ.PC 2m7#1). This clone was transfected under standard conditions (Ausubel et al., 1994, Vol. 1 and 2 "Current Protocols in Molecular Biology" John Wiley & Sons, Inc.) with a eukaryotic vector (pEF-BOS; promoter originates from the human EF-1 alpha gene, selection marker: puromycin; 15 Mizushima and Nagata, 1990, Nucleic Acids Res. 18: 5322), which contains the cDNA of the human B7.1 gene (Selvakumar et al., 1992, Immunogenetics 36: 175-181). A clone MZ-PC2 m7#1 B7.1#3 was selected and cultured in T150 cell culture flasks. The nutrient medium used was RPMI 1640 (Gibco plus 4g/L glucose) containing 10% heat-inactivated foetal calf serum and 2 mM of L-glutamine. Every 3 to 4 days the cells were cleaved for 20 propagation by trypsinisation at 1:5. After about 80% confluence had been achieved 4 ml of a trypsin solution (containing per litre: 8g NaCl, 0.2g KCl, 1.13g Na₂HPO₄-anhydrous, 0.2g KH₂PO₄, 100ml 2.5% trypsin solution, 1g of EDTA-Na-salt; pH 7.2-7.4) were added to each T150 cell culture flask to harvest the cells. In all, 2x10⁷ cells were used to isolate the RNA according to the manufacturer's instructions (RNeasy Minikit, QIAgen). Starting 25 from about 100 µg of total-RNA the manufacturer's instructions were followed in order to isolate polyA⁺ RNA using the Oligotex Kit (QIAgen). Then starting with about 0.5 mg of polyA⁺ RNA the cDNA synthesis was carried out according to the manufacturer's instructions (Clontech Marathon Protokoll).

30 Example 2

Representational Difference Analysis (RDA) of the pancreatic carcinoma cell line MZ.PC2 m7#1 B7.1#3 versus a pool of 11 normal tissues

Starting from about 0.5 µg of poly-A(+) of the pancreatic tumour cell line MZ.PC2 m7#1 B7.1#3 and a pool of 2.5 µg of poly-A(+) RNA from 11 normal tissues (Clontech) - bone marrow, heart, kidney, liver, lung, pancreas, skeletal muscle, spleen, thymus, small intestine and stomach - RDA was carried out (Diatchenko et al. 1996, Proc. Natl. Acad. Sci. U.S.A. 93: 6025-6030, ; Hubank and Schatz, 1994, Nucleic. Acids. Res. 22: 5640-5648) using the PCR-select™ Kit (Clontech, Palo Alto) in accordance with the manufacturer's instructions: RNA from the pancreatic tumour cell line was used as the "tester" and RNA from the normal tissue pool was used as the "driver" in accordance with the manufacturer's instructions. In contrast to the original procedure, after the synthesis of double-stranded cDNA using oligo-dT, the cDNA was cut with 6 restriction enzymes: *EcoRV*, *NaeI*, *NruI*, *ScaI* (Promega), *SspI*, *StuI* (TaKaRa) in Promega buffer A for 2 hours at 37°C and, after increasing the NaCl concentration to 150mM, for a further 2 hours at 37°C. The use of this mixture of 6 different restriction enzymes made it possible to generate cDNA fragments about 800 bp long, which were used for the representational difference analysis.

Equal parts of tester cDNA were ligated with either adaptor A or B and then hybridised separately with an excess of driver-cDNA at 68°C. Then the two mixtures were combined and subjected to a second hybridisation with fresh denatured driver cDNA. The concentrated tester-specific cDNAs were then amplified exponentially by PCR with primers from the kit specific for the adaptor A or B with a 2 minute elongation time at 72°C, over 27 cycles (10'' at 94°C, 30'' at 66°C, 2' at 72°C). For further concentration, one aliquot from this reaction was subjected to a second PCR with specific nested primers from the kit with 2 minutes' elongation time at 72°C, 10 cycles (10'' at 94°C, 30'' at 66°C, 2' at 72°C). The product resulting from this reaction was ligated into 3 different, individually modified pCR3.1(+) vectors (Invitrogen): vector (1.ORF), vector (2.ORF) and vector (3.ORF) (Fig. 5: CMV Cytomegalovirus; BGH Bovine Growth Hormone; ORF Open Reading Frame) and then transformed into competent *E. coli* (OneShot™, Invitrogen). These vectors allow expression in eukaryotic cells in 3 different reading frames.

To construct the 3 vectors, the pCR3.1(+)-vector (Invitrogen) was cut with *NheI* and *HindIII* (Promega) and ligated with one dsDNA oligomer which was produced by annealing two ssDNA oligomers (SEQ ID NO:4 and 5; vector ORF1) or (SEQ ID NO:6 and 7; vector ORF2) or (SEQ ID NO:8 and 9; vector ORF3), using standard methods (e.g.

Ausubel et al., 1994; Sambrook et al. 1989 ColdSpring Harbor Laboratory Press). The 3 vector types have a start codon and a cloning site for expression in a reading frame which is different from the other two vectors.

5 The transformation of competent *E. coli* (OneShot™, Invitrogen) carried out in three batches (vector 1.ORF, 2.ORF and 3.ORF) with the cDNA of the subtractive cDNA library produced about 9600 clones. These were examined by PCR analysis for the presence and length of the insert cDNA. The following method was used: the 9600 clones were cultivated in 96-well blocks in LB-Amp medium for 48 h at 37°C. Then 5 µl aliquots of
10 the *E. coli* suspensions were heated to 100°C in 500 µl of TE buffer for 10 minutes and 1.5 µl thereof were used as the basis for a PCR in which the insert of the vector was amplified with flanking primers (SEQ ID NO:10 and 11) over 35 cycles (1' at 94°C, 1' at 55°C, 2' at 72°C). The PCR products were revealed by agarose-gel electrophoresis and ethidium bromide staining. The bacterial cultures remaining were stored as glycerol stock cultures
15 at -80°C.

A cDNA subtraction library of 3450 individual clones was obtained in the form of *E. coli* glycerol stock cultures, the insert length of which was known from agarose gel electrophoresis. As expected, the inserted cDNA fragments were shown to have an
20 average length of about 800 bp.

Example 3

25 DNA Sequencing and Annotation of clones of the subtractive cDNA library of the pancreatic tumour cell line MZ.PC2 m7#1 B7.1#3

The plasmid-DNA from 50 clones randomly selected from the subtractive cDNA library were isolated in accordance with the manufacturer's instructions (QIAGEN) and sequenced by the Sanger method on an ABI-Prism apparatus. The sequences thus found were
30 annotated by BLAST-Search (National Center for Biotechnology Information) and subjected to EST data bank comparisons. This made it possible to identify 38 known and 12 unknown genes. For the latter there were only EST entries. For the 12 unknown genes the expression profile was estimated: the starting tissue for the corresponding cDNA library was checked for all the ESTs in data banks having greater than 95% identity
35 (BLAST) with the experimentally determined sequence. They were subdivided into

i) critical normal tissue, ii) foetal, "non-essential" and immunoprivileged tissue and
 iii) tumours and tumour cell lines. On the basis of this "virtual mRNA profile" 4 clones
 (R2, R8, R11 and R12) were selected for further experimental analysis.

5 Example 4

Transcriptional analysis of the candidate clones in tumour and normal tissue

Between 2 and 5 µg of total RNA from tumour or normal tissues were reverse transcribed
 10 using SuperScriptII (GibcoBRL) or AMV-RT (Promega) in accordance with the
 manufacturer's recommendations. For each individual RNA probe a second test was
 carried out without reverse transcriptase as a control for contamination by chromosomal
 DNA. The quality and quantity of the cDNAs was checked by PCR with β-actin-specific
 primers (SEQ ID NO: 14 and 15) and GAPDH specific primers (SEQ ID NO:16 and 17)
 15 after 30 and 35 cycles (1' at 95°C, 1' at 55°C, 1' at 72°C). The 4 candidate genes were
 analysed analogously with specific primers. The PCR products were detected by agarose
 gel electrophoresis and ethidium bromide staining. A candidate which was designated
 "R11" exhibited a relatively specific tumour/testis transcription profile, after 35 cycles with
 R11-specific primers (SEQ ID NO:12 and 13); the semiquantitative RT-PCR of RNA from
 20 carcinoma of the breast, adenocarcinoma of the lung, plate epithelial carcinoma of the
 lung, carcinoma of the kidney, colon carcinoma, heart, lung, liver, kidney, colon, spleen
 and testis is shown in Fig. 1. Another qualitative PCR of cDNA from the tissue of 3
 human patients with tumours of the pancreas using the same R11-specific primers (SEQ ID
 NO:12 and 13) showed expression in human pancreatic tumours (Fig. 2). Moreover, an
 25 additional qualitative PCR of cDNA from various tumour cell lines from human lung (LC
 6, 16), gall bladder (GB 1) and pancreatic tumours (PC 1, 2) and two melanomas (Mel 2,
 7) was carried out with the same R11-specific primers (SEQ ID NO:12 and 13), which
 showed clear expression in all the tumour cell lines (Fig. 4). In this analysis the Perkin
 Elmer method (GeneAmp RNA PCR Kit, #N808-0017) was used (RT reaction: (1x)
 30 15'/42°C - 5'/99°C - 5'/4°C; PCR reaction: (35x) 2'/95°C - 1'/95°C - 1'/60°C and (1x)
 7'/72°C - 4°C (Fig. 4). As described above, the PCR products were revealed by agarose
 gel electrophoresis and ethidium bromide staining. A 1kb size marker made by Gibco
 BRL was used as the size marker.

Example 5

Transcription profile of R11 in normal tissues

5 For Northern Blot analysis, Human Multiple Tissue Northern Blots (Clontech, Palo Alto and Invitrogen) were hybridised for 2 h at 68° with the roughly 1000 bp long R11 PCR product labelled with [α -³²P]dCTP (NEN, Boston). Visualisation was carried out by standard autoradiography (Hyperfilm, Amersham). Fig. 3 shows the results of this analysis: from 19 normal tissues (pancreas, adrenal medulla, thyroid, adrenal cortex, testes, 10 thymus, small intestine, stomach, brain, heart, skeletal muscle, colon, spleen, kidney, liver, placenta, lung, leukocytes). For R11, a prominent transcript 7.5 kb long is found in the placenta, adrenal medulla, adrenal cortex and in the testis. A very weak band of 7.5 kb can also be detected in the brain. Since all these normal tissues have immunoprivileged status (Streilein, 1995, Science 270: 1158-1159) an attack by CTL can be ruled out in any 15 immunotherapy based on this antigen.

Other transcripts of 3.8 kb and 2.3 kb, which might possibly be splice variants of the 7.5 kb transcript or might be derived from a homologous gene, were identified in the adrenal medulla and adrenal cortex (Fig. 3).

Example 6

Cloning the R11 cDNA

25 The following procedure was used to clone the human R11 cDNA: a BLAST search identified a fragment AF038197 and a plurality of ESTs, such as for example N42343, W69539, H82474, H51766, N28313, overlapping with the R11 "original sequence" (796 bp) obtained in Example 3 by sequencing. Starting from the sequence AF038197 a contig overlapping with the clone R11 was found with the EstExtractor on TigemNet 30 (<http://gcg.tigem.it/cgi-bin/uniestass.pl>, Banfi et al., 1996, Nature Genetics 13:167-174). The overlapping of the contig and the "original sequence" of the 796 bp long R11 was verified by PCR amplification with an R11 "original sequence"- specific primer and a primer located on the contig (SEQ ID NO: 18 and 19) from a SuperScript™ Human Testis cDNA Library (GibcoBRL) and subsequent sequencing. By means of a PCR with an R11- 35 specific primer (SEQ ID NO:20) and a vector-specific primer (SEQ ID NO:21), other

fragments belonging to R11 were amplified from the SuperScript™ Human Testis cDNA Library using the Advantage cDNA PCR Kit (Clontech) and following the standard procedure described therein. Knowledge of these new sequences in turn made it possible to carry out further PCRs with R11-specific primers (SEQ ID NO: 22 and 23) and the
 5 vector-specific primer (SEQ ID NO:21).

To extend the R11-cDNA further, a human testis Rapid-Screening cDNA Library panel (OriGene Technologies, Inc) was screened with primers specific to R11 (SEQ ID NO: 24 and 25) under the standard PCR conditions specified by the manufacturer. From the
 10 positive wells, one aliquot was amplified as a template for a PCR with an R11 specific primer and a primer specific to the vector (SEQ ID NO: 26 and 27) using the Advantage cDNA PCR Kit (Clontech) and following the standard procedure described therein.

For the sequence analysis, aliquots of the PCR preparations were ligated directly into the
 15 pCR2.1 vector (Invitrogen) and then transformed into competent *E. coli* (OneShot™, Invitrogen) and sequenced as described in Example 3.

Starting from these newly identified sequences, 5'-sequence regions located higher upstream could be cloned from a SuperScript™ Human Testis cDNA Library (GibcoBRL) with the following additional oligonucleotide primers specific for R11 (SEQ ID NOs:28 to
 20 43). The primers were used with a plasmid-specific primer (SEQ ID NO:21) described hereinbefore or combined with one another for the PCR cloning using the Advantage cDNA PCR Kit (Clontech).

25 The cloned region of the R11-cDNA has 6582 bp, whilst the presence of a PolyA tail at the 3'-end of the sequence is an indication of the completeness of the cDNA in this region. Two separate continuous reading frames were identified. The first reading frame at the 5' end (R11-ORF-1; SEQ ID NO:2) is represented by the start codon at position 218 and the stop codon (TAG) at position 1421 in SEQ ID NO:1. There are no data bank entries of
 30 known genes for this gene. Analysis of the protein profile (<http://www.expasy.ch/prosite>, Hofmann et al., 1999, Nucleic Acid Res. 27:215-219) yielded a reference to three possible N-glycosylation sites (position # 62-65, 76-79 and 117-120 in SEQ ID NO: 2), a cAMP- and cGMP-dependent protein kinase phosphorylation site (position # 11-14 in SEQ ID NO.2), as well as 7 possible PKC-phosphorylation sites (position # 9-11, 14-16, 78-80,

119-121, 183-185, 202-204 and 210-212 in SEQ ID NO:2) and 6 possible casein kinase II phosphorylation sites (position # 119-122, 127-130, 183-186, 256-259, 295-298 and 358-361 in SEQ ID NO:2). The zinc finger motif (zf-CCHC; E=0,11, Pfam-A HMM) from position # 371 to position # 384 in SEQ ID NO:2 (CLYCGTGGHYADNC) should be of particular interest for predicting a possible function of R11-ORF-1. It is known that members of the protein family which have these motifs do not have any insertions or deletions in the motif itself; this is also true of the R11-ORF-1 protein. Although no typical SH3 binding sequence can be found, it is certainly conceivable that the P-rich region (position # 36-56 in SEQ ID NO:2) could interact with an SH3 domain. By using the COILS algorithm, a coiled-coil structure can be predicted with more than 99% probability for the amino acid groups in the region of position # 80 to about 125. On the basis of these two domains, the zinc finger motif and the coiled coil domain, it can be concluded that R11-ORF-1 is possibly a transcription factor the oligomerisation of which is controlled via these two domains.

In the second open reading frame, R11-ORF-2, which is defined by a start codon at position # 1498 and a stop codon (TAA) at 2569, in addition to the two obvious proline-rich sections (position # 128-141 and 330-351 in SEQ ID NO: 2), potential motifs for two N-glycosylation sites (104-107 and 251-254), one protein kinase C phosphorylation site (108-110), five casein kinase II phosphorylation sites (99-102, 165-168, 198-201, 200-203 and 274-277) and a region resembling the active centre of eukaryotic and viral aspartate proteinases (16-27). The clear homology of the first 280 amino acids of R11-ORF-2 with the retroviral pol polyprotein is particularly remarkable. In the C-terminus, by contrast, no homologies could be discovered. Amino acids from position #9 to 277 clearly align in blastp with the Fugu pol polyprotein (position # 104-365; $2e^{-22}$). The aspartate protease pattern # 16-27 mentioned above comprises the active nucleophil Asp (#19) of the active centre of the protease of the pol region; position #215 to #277 corresponding to part of the reverse transcriptase domain. The protein derived from R11-ORF-2 is therefore a possible retrotransposon.

Example 7

Potential MHC-binding peptides in the regions coding for the two reading frames of R11, R11-ORF-1 and R11-ORF-2

- Potential peptide epitopes within the two reading frames of R11 according to SEQ ID NO:2 or 3) were carried out using the algorithms described by Parker et al., 1994, J. Immunol. 152: 163 on the basis of known motifs (Rammensee et al., 1995, Immunogenetics 41:178-228). 9-mer candidate peptides have been identified for the most important HLA-types, especially for HLA-A1, -A*0201, -A3, -B7, -B14 and -B*4403, which can be expected to bind to the corresponding HLA molecules and thus constitute immunogenic CTL-epitopes; the peptides discovered are listed in Table 1 (R11-ORF-1) and Table 2 (R11-ORF-2). By analogous methods, other potential peptide epitopes may be found for other HLA types or 8-, 9-, or 10-mer peptides.

Table 1

Immunogenic peptide candidates of R11-ORF-1 (401 amino acids)

Starting position in SEQ ID NO:2	Sequence	HLA
35	Ser Pro Pro Thr Pro Thr Val Thr Leu (SEQ ID NO:88)	HLA-B7
85	Leu Ser Glu Glu Ile Asn Asn Leu Arg (SEQ ID NO:89)	HLA-A1
112	Lys Leu Thr Glu Glu Asn Thr Thr Leu (SEQ ID NO:90)	HLA-A*0201
113	Leu Thr Glu Glu Asn Thr Thr Leu Arg (SEQ ID NO:91)	HLA-A1
135	Ile Glu Leu Arg Gly Ala Ala Ala Ala (SEQ ID NO:92)	HLA-B*4403
172	Phe Met Ala Gln Cys Gln Ile Phe Met (SEQ ID NO:93)	HLA-A*0201
199	Ser Met Met Thr Gly Arg Ala Ala Arg (SEQ ID NO:94)	HLA-A3
205	Ala Ala Arg Trp Ala Ser Ala Lys Leu (SEQ ID NO:95)	HLA-B7

211	Ala Lys Leu Glu Arg Ser His Tyr Leu (SEQ ID NO:96)	HLA-B14
252	Gln Gly Met Gly Ser Val Ile Asp Tyr (SEQ ID NO:97)	HLA-B*4403
274	Asn Glu Pro Ala Leu Ile Asp Gln Tyr (SEQ ID NO:98)	HLA-B*4403
315	Arg Arg Leu Ala Arg Ala Ala Ala Ala (SEQ ID NO:99)	HLA-B14
325	Lys Pro Arg Ser Pro Pro Arg Ala Leu (SEQ ID NO:100)	HLA-B7
354	Arg Met Arg Leu Thr Gln Glu Glu Lys (SEQ ID NO:101)	HLA-A3
346	Pro Thr Glu Pro Val Gly Gly Ala Arg (SEQ ID NO:102)	HLA-A1

Table 2

Immunogenic peptide candidates of R11-ORF-2 (357 amino acids)

Starting position in SEQ ID NO:3	Sequence	HLA
1	Met Leu Gln Ile His Leu Pro Gly Arg (SEQ ID NO:44)	HLA-A3
4	Ile His Leu Pro Gly Arg His Thr Leu (SEQ ID NO:45)	HLA-A*0201, HLA-B14
5	His Leu Pro Gly Arg His Thr Leu Phe (SEQ ID NO:46)	HLA-A3
31	Tyr Val Ala Gln Asn Gly Ile Pro Leu (SEQ ID NO:47)	HLA-B7
39	Leu Arg Ile Lys Asp Trp Pro Ile Leu (SEQ ID NO:48)	HLA-B14
46	Ile Leu Val Glu Ala Ile Asp Gly Arg (SEQ ID NO:49)	HLA-A3
53	Gly Arg Pro Ile Ala Ser Gly Pro Val (SEQ ID NO:50)	HLA-B14
64	Glu Thr His Asp Leu Ile Val Asp Leu (SEQ ID NO:51)	HLA-B14
71	Asp Leu Gly Asp His Arg Glu Val Leu (SEQ ID NO:52)	HLA-A*0201, HLA-B7, HLA-B14
73	Gly Asp His Arg Glu Val Leu Ser Phe (SEQ ID NO:53)	HLA-B*4403
85	Gln Ser Pro Phe Phe Pro Val Val Leu (SEQ ID NO:54)	HLA-B7

92	Val Leu Gly Pro Arg Trp Leu Ser Ala (SEQ ID NO:55)	HLA-A*0201
97	Trp Leu Ser Ala His Asp Pro Asn Ile (SEQ ID NO:56)	HLA-A*0201
110	Arg Ser Ile Val Phe Asp Ser Glu Tyr (SEQ ID NO:57)	HLA-B*4403
112	Ile Val Phe Asp Ser Glu Tyr Cys Arg (SEQ ID NO:58)	HLA-A3
134	Pro Pro Pro Ala Pro Gln Pro Pro Leu (SEQ ID NO:59)	HLA-B7
141	Pro Leu Tyr Tyr Pro Val Asp Gly Tyr (SEQ ID NO:60)	HLA-A3
150	Arg Val Tyr Gln Pro Val Arg Tyr Tyr (SEQ ID NO:61)	HLA-A3
152	Tyr Gln Pro Val Arg Tyr Tyr Tyr Val (SEQ ID NO:62)	HLA-A*0201
155	Val Arg Tyr Tyr Tyr Val Gln Asn Val (SEQ ID NO:63)	HLA-B14
159	Tyr Val Gln Asn Val Tyr Thr Pro Val (SEQ ID NO:64)	HLA-A*0201
169	Gly His Val Tyr Pro Asp His Arg Leu (SEQ ID NO:65)	HLA-B14
177	Leu Val Asp Pro His Ile Glu Met Ile (SEQ ID NO:66)	HLA-A*0201, HLA-A1
183	Glu Met Ile Pro Gly Ala His Ser Ile (SEQ ID NO:67)	HLA-A*0201
189	His Ser Ile Pro Ser Gly His Val Tyr (SEQ ID NO:68)	HLA-A1, HLA-B*4403
191	Ile Pro Ser Gly His Val Tyr Ser Leu (SEQ ID NO:69)	HLA-B7, HLA-A*0201

198	Ser Leu Ser Glu Pro Glu Met Ala Ala (SEQ ID NO:70)	HLA-A*0201
199	Leu Ser Glu Pro Glu Met Ala Ala Leu (SEQ ID NO:71)	HLA-A1
202	Pro Glu Met Ala Ala Leu Arg Asp Phe (SEQ ID NO:72)	HLA-B*4403
203	Glu Met Ala Ala Leu Arg Asp Phe Val (SEQ ID NO:73)	HLA-A*0201
206	Ala Leu Arg Asp Phe Val Ala Arg Asn (SEQ ID NO:74)	HLA-A*0201
211	Val Ala Arg Asn Lys Asp Gly Leu (SEQ ID NO:75)	HLA-B7
223	Thr Ile Ala Pro Asn Gly Ala Gln Val (SEQ ID NO:76)	HLA-A*0201
224	Ile Ala Pro Asn Gly Ala Gln Val Leu (SEQ ID NO:77)	HLA-B7
231	Val Leu Gln Val Lys Arg Gly Trp Lys (SEQ ID NO:78)	HLA-A3
232	Leu Gln Val Lys Arg Gly Trp Lys Leu (SEQ ID NO:79)	HLA-A*0201, HLA-B14
258	Tyr Pro Arg Leu Ser Ile Pro Asn Leu (SEQ ID NO:80)	HLA*7
266	Glu Asp Gln Ala His Leu Ala Thr Tyr (SEQ ID NO:81)	HLA-B*4403
270	His Leu Ala Thr Tyr Thr Glu Phe Val (SEQ ID NO:82)	HLA-A*0201
306	Gly Arg Asp Gly Gln Gly Arg Ser Leu (SEQ ID NO:83)	HLA-B14
307	Arg Asp Gly Gln Gly Arg Ser Leu Tyr (SEQ ID NO:84)	HLA-B*4403

308	Asp Gly Gln Gly Arg Ser Leu Tyr Val (SEQ ID NO:85)	HLA-B14
313	Ser Leu Tyr Val Pro Val Met Ile Thr (SEQ ID NO:86)	HLA-A*0201, HLA-A3
320	Ile Thr Trp Asn Pro His Trp Tyr Arg (SEQ ID NO:87)	HLA-A3

The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention. Indeed various modifications of the invention in addition to those shown and described herein will become
5 apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

All publications and patent applications cited herein are incorporated by reference in their entireties.

What is claimed is:

1. A tumour-associated antigen (TAA) comprising the amino acid sequence of SEQ ID NO:2.
- 5 2. A tumour-associated antigen (TAA) designated R11-ORF-1 consisting of the amino acid sequence of SEQ ID NO:2.
3. An immunogenic protein fragment derived from the TAA of claim 1.
- 10 4. The TAA of claim 1 wherein the TAA induces or augments a humoral immune response.
5. The immunogenic protein fragment of claim 3 wherein the protein fragment induces or
15 augments a humoral immune response.
6. The TAA of claim 1 wherein the TAA or a breakdown product thereof is presented by an MHC molecule and induces or augments a cellular immune response.
- 20 7. The immunogenic protein fragment according to claim 3 wherein the protein fragment or a breakdown product thereof is presented by an MHC molecule and induces or augments a cellular immune response.
8. The immunogenic protein fragment of claim 7, wherein the amino acid sequence of the
25 protein fragment is selected from the group consisting of SEQ ID NOs:88 to 102.
9. A derivative of the TAA of claim 1 wherein the derivative displays one or more functional activities of the TAA.

10. A method for treating or preventing cancer in an individual comprising administering to the individual the TAA of claim 1, wherein the TAA induces or augments an immune response in the individual against tumor cells which express R11.
- 5 11. A method for treating or preventing cancer in an individual comprising administering to the individual the immunogenic protein fragment of claim 3, wherein the protein fragment induces or augments an immune response in the individual against tumor cells which express R11.
- 10 12. A method for treating or preventing cancer in an individual comprising:
- (a) incubating cytotoxic T lymphocyte precursor cells (CTLs) with antigen presenting cells and the tumor-associated antigen (TAA) of claim 1 wherein the CTLs become activated; and
 - (b) administering the activated CTLs to an individual,
- 15 wherein the TAA induces or augments an immune response in the individual against tumor cells which express R11.
13. A method for treating or preventing cancer in an individual comprising:
- (a) incubating cytotoxic T lymphocyte precursor cells (CTLs) with antigen presenting cells and the immunogenic protein fragment of claim 3 wherein the CTLs become activated; and
 - (b) administering the activated CTLs to an individual,
- 20 wherein the protein fragment induces or augments an immune response in the individual against tumor cells which express R11.
- 25 14. The method of any one of claims 10 to 13 wherein the individual is a human.
15. The method of any one of claims 10 to 13 wherein the cancer is selected from the group consisting of kidney carcinoma, breast carcinoma and pancreatic carcinoma.
- 30

16. The method of claim 10 or 11 wherein the tumor-associated antigen or the immunogenic protein fragment is administered parenterally, topically, orally, or locally.
- 5 17. A pharmaceutical composition comprising the tumor-associated antigen of claim 1; and a pharmaceutically acceptable carrier.
18. A tumour-associated antigen (TAA) designated R11-ORF-2 consisting of the amino acid sequence of SEQ ID NO:3.
- 10 19. An immunogenic protein fragment derived from the TAA of claim 18.
20. The TAA of claim 18 wherein the TAA induces or augments a humoral immune response.
- 15 21. The immunogenic protein fragment of claim 19 wherein the protein fragment induces or augments a humoral immune response.
22. The TAA of claim 18 wherein the TAA or a breakdown product thereof is presented by an MHC molecule and induces or augments a cellular immune response.
- 20 23. The immunogenic protein fragment according to claim 19 wherein the protein fragment or a breakdown product thereof is presented by an MHC molecule and induces or augments a cellular immune response.
- 25 24. The immunogenic protein fragment of claim 23, wherein the amino acid sequence of the protein fragment is selected from the group consisting of SEQ ID NOs:44 to 87.
- 30 25. A derivative of the TAA of claim 18 wherein the derivative displays one or more functional activities of the TAA.

26. A method for treating or preventing cancer in an individual comprising administering to the individual the TAA of claim 18, wherein the TAA induces or augments an immune response in the individual against tumor cells which express R11.

5 27. A method for treating or preventing cancer in an individual comprising administering to the individual the immunogenic protein fragment of claim 19, wherein the protein fragment induces or augments an immune response in the individual against tumor cells which express R11-ORF.

10 28. A method for treating or preventing cancer in an individual comprising:
 (c) incubating cytotoxic T lymphocyte precursor cells (CTLs) with antigen presenting cells and the tumor-associated antigen (TAA) of claim 18 wherein the CTLs become activated; and
 (d) administering the activated CTLs to an individual,
 15 wherein the TAA induces or augments an immune response in the individual against tumor cells which express R11.

29. A method for treating or preventing cancer in an individual comprising:
 (c) incubating cytotoxic T lymphocyte precursor cells (CTLs) with antigen
 20 presenting cells and the immunogenic protein fragment of claim 19 wherein the CTLs become activated; and
 (d) administering the activated CTLs to an individual,
 wherein the protein fragment induces or augments an immune response in the individual against tumor cells which express R11.

25 30. The method of any one of claims 26 to 29 wherein the individual is a human.

31. The method of any one of claims 26 to 29 wherein the cancer is selected from the group consisting of kidney carcinoma, breast carcinoma and pancreatic carcinoma.

30

32. The method of claim 26 or 27 wherein the tumor-associated antigen or the immunogenic protein fragment is administered parenterally, topically, orally, or locally.
- 5 33. A pharmaceutical composition comprising the tumor-associated antigen of claim 1 or claim 18; and a pharmaceutically acceptable carrier.
34. A pharmaceutical composition comprising one or more immunogenic protein fragments derived from the tumor-associated antigen of claim 1 or claim 18; and a
10 pharmaceutically acceptable carrier.
35. The pharmaceutical composition according to claim 34 further comprising one or more isolated protein fragments derived from a tumor-associated antigen which is not R11.
- 15 36. The pharmaceutical composition according to claim 34 wherein the protein fragments bind to at least two different HLA types.
37. The pharmaceutical composition according to claim 35 wherein the protein fragments bind to at least two different HLA types.
20
38. An isolated nucleic acid comprising a nucleotide sequence encoding a protein or protein fragment, which protein or protein fragment displays one or more functional activities of the TAA of claim 1.
- 25 39. An isolated nucleic acid comprising a nucleotide sequence encoding the tumor-associated antigen of claim 2 or claim 18.
40. An isolated nucleic acid comprising a nucleotide sequence encoding the immunogenic protein fragment of claim 3 or claim 19.
30

41. An isolated nucleic acid consisting of the nucleotide sequence of SEQ ID NO:1, or a degenerate thereof.
42. An isolated nucleic acid that hybridizes to a nucleotide sequence of SEQ ID NO:1,
5 under stringent conditions.
43. A recombinant DNA molecule comprising the nucleic acid of any of claims 38 to 42.
44. A recombinant cell containing a recombinant nucleic acid vector comprising the
10 nucleic acid of any of claims 38 to 42.
45. A method of producing a protein comprising:
- (a) growing the recombinant cell of claim 44, such that the protein is expressed by the cell; and
 - 15 (b) recovering the expressed protein.
46. A method for treating or preventing cancer comprising administering to an individual a therapeutic or prophylactic amount of the nucleic acid of any of claims 38 to 42, wherein the nucleic acid induces or augments an immune response in the individual
20 against tumor cells which express R11.
47. A pharmaceutical composition comprising a nucleic acid of any of claims 38 to 42, and a pharmaceutically acceptable carrier.
- 25 48. An antibody which specifically binds the tumor-associated antigen of any one of claims 1, 2 or 18.
49. An antibody which specifically binds the immunogenic protein fragment of any one of claims 3, 8, 19 or 24.
- 30 50. The antibody of claim 48 which is a monoclonal antibody.

51. The antibody of claim 49 which is a monoclonal antibody.

52. A method for treating or preventing cancer comprising administering the antibody of claim 48, wherein the antibody induces or augments an immune response in the individual against tumor cells which express R11.

53. A method for treating or preventing cancer comprising administering the antibody of claim 49, wherein the antibody induces or augments an immune response in the individual against tumor cells which express R11.

54. A method of diagnosing or screening for cancer comprising detecting or measuring in a sample derived from an individual the level of the TAA of any one of claims 1, 2 or 18, wherein an increase in the level of the TAA or nucleic acid encoding the TAA indicates the presence of cancer associated with the expression of the TAA.

55. A method for detecting tumor cells that express the TAA of any one of claims 1, 2 or 18 in an individual, comprising imaging said individual after administration to said individual of a labeled molecule which specifically binds to the TAA, wherein detection of the labeled molecule above background indicates the presence of a tumor that expresses the TAA.

56. The method of claim 55 wherein the labeled molecule is an antibody or a portion of said antibody containing the binding domain thereof which specifically binds a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:3.

57. The method of claim 55 wherein the labeled molecule is a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:3.

58. A kit comprising in one or more containers a molecule selected from the group consisting of the TAA of any one of claims 1, 2 or 18 or a fragment or a derivative

thereof; and, in another container, an antibody that specifically binds the TAA or the fragment or the derivative.

11/11/2011 11:11:11 AM

ABSTRACT

5 Tumour-associated antigens, immunogenic peptides derived therefrom and DNA molecules coding therefor, and the use thereof in the immunotherapy of cancers.

Fig. 1A

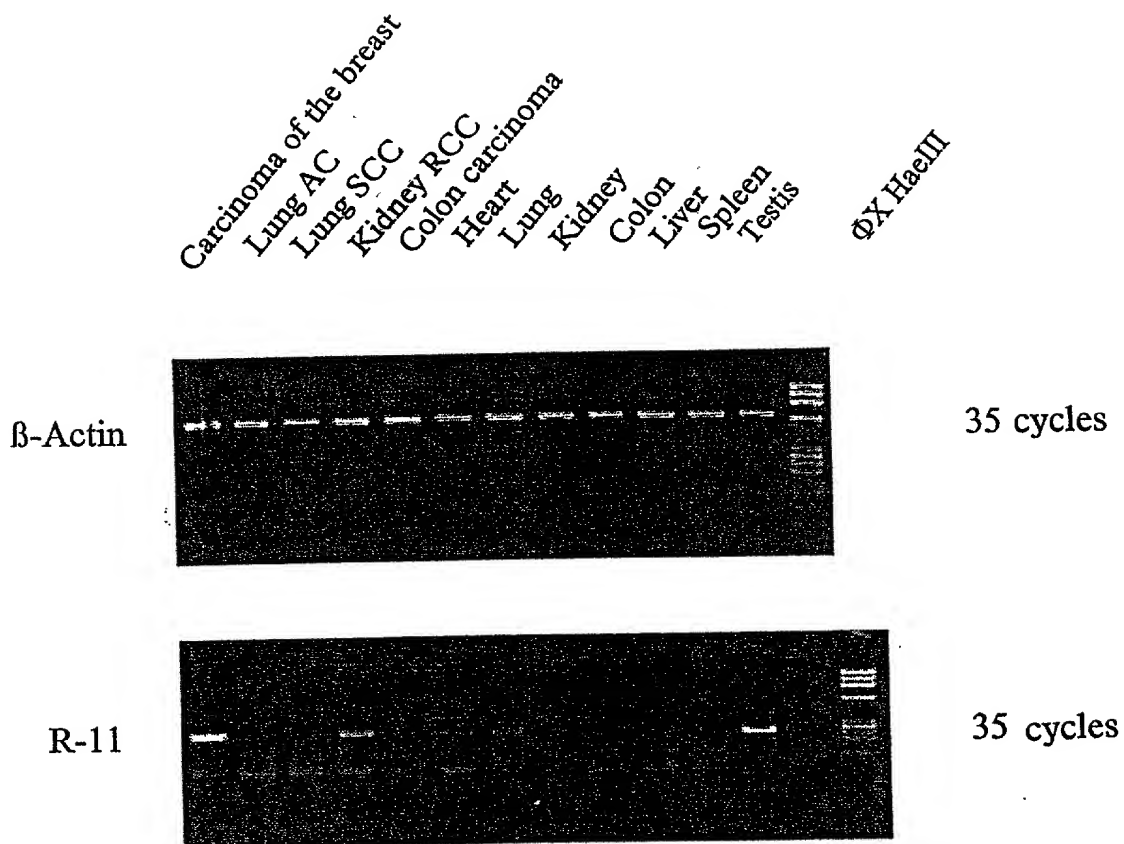


Fig. 1B

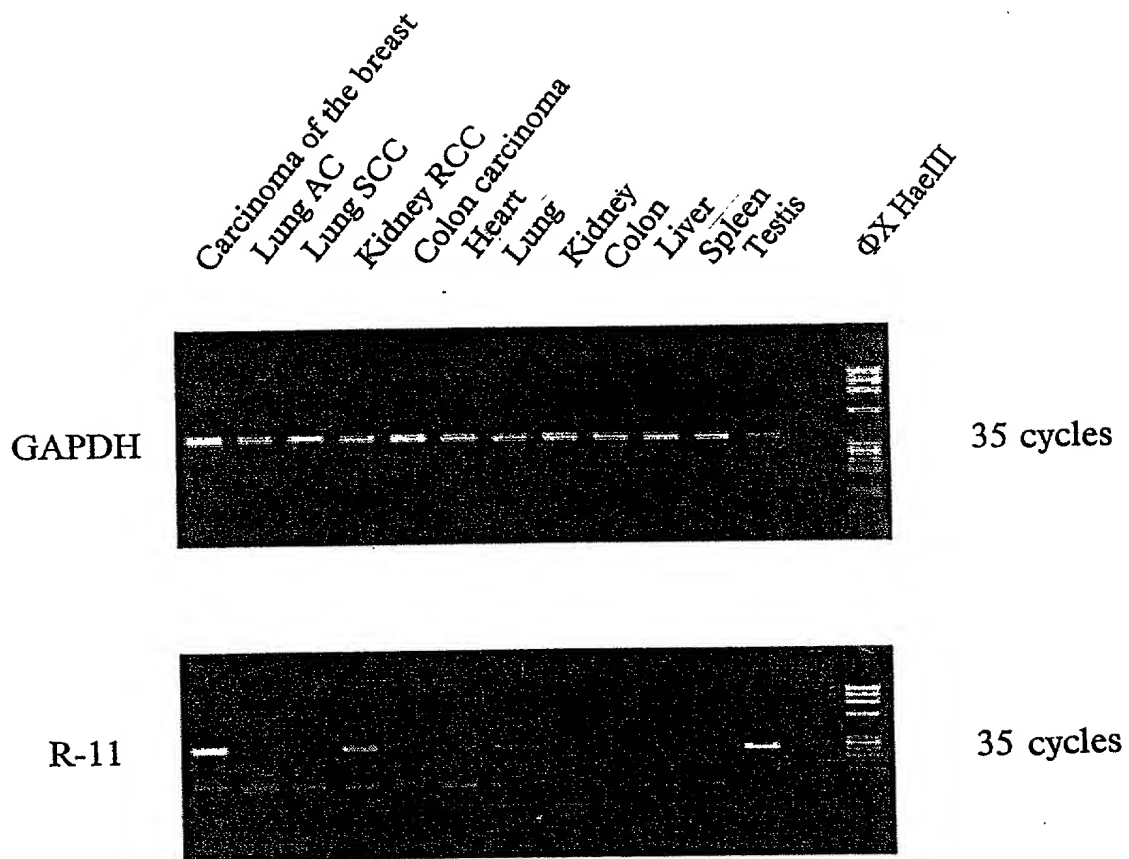


Fig. 2

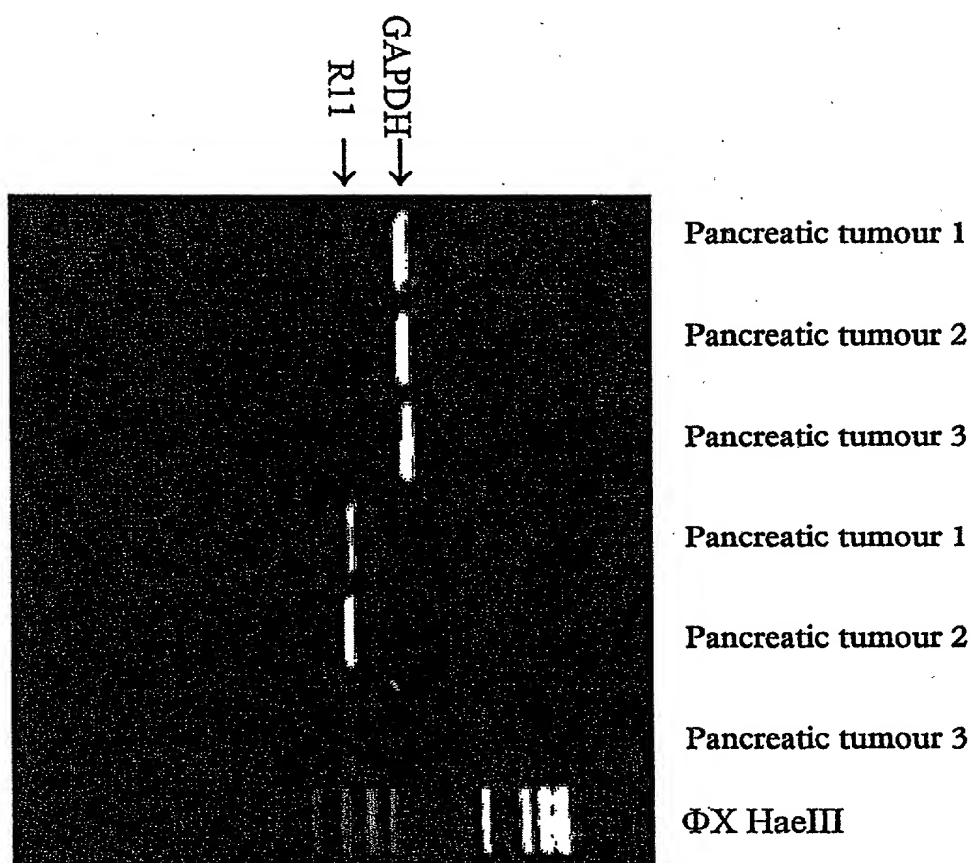


Fig. 3

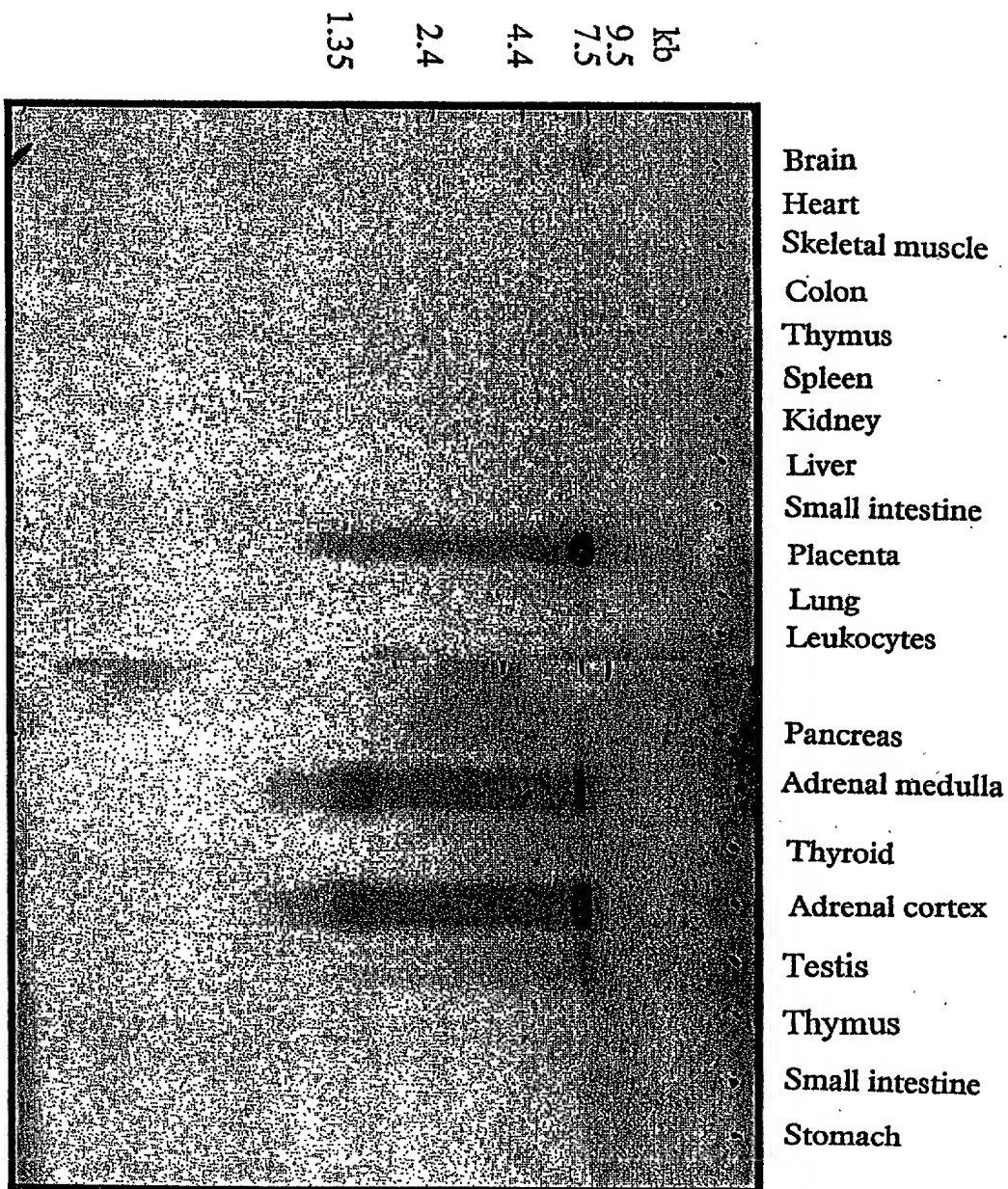
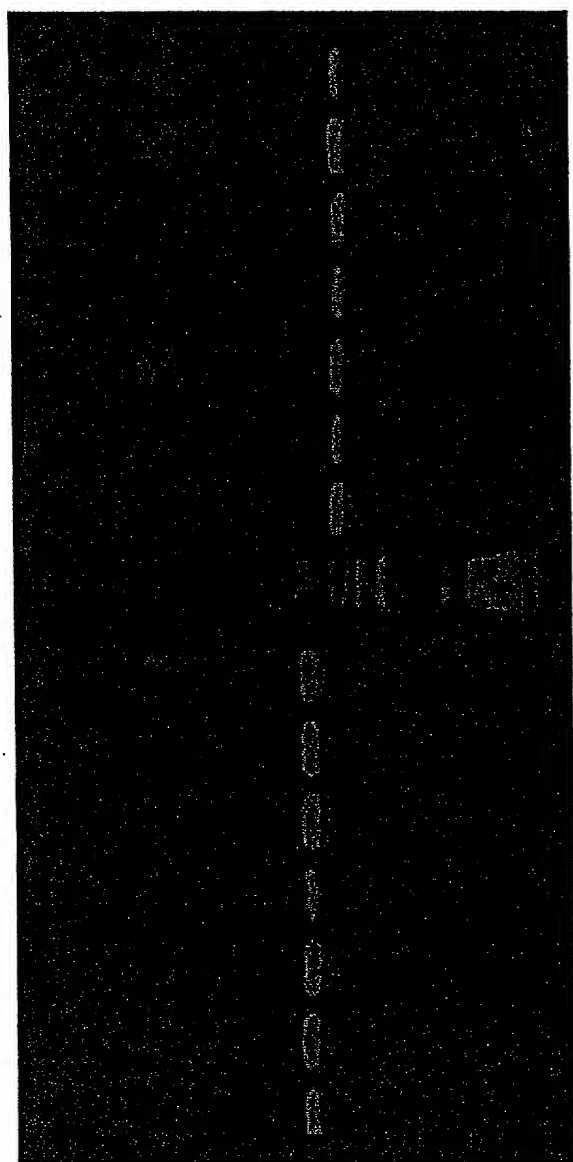


Fig. 4

(35 cycles)



LC 6

LC 16

GB 1

Mel 2

Mel 7

PC 1

PC 2

Marker

LC 6

LC 16

GB 1

Mel 2

Mel 7

PC 1

PC 2

Please type a plus sign (+) inside this box → ☐

PTO/SB/01 (12-97)

Approved for use through 9/30/00. OMB 0651-0032
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION (37 CFR 1.63) <input type="checkbox"/> Declaration Submitted with Initial Filing OR <input checked="" type="checkbox"/> Declaration Submitted after Initial Filing (surcharge (37 CFR 1.16 (e)) required)	Attorney Docket Number	12/211
	First Named Inventor	Konopitzky, Renate
	COMPLETE IF KNOWN	
	Application Number	/ To be assigned
	Filing Date	August 3, 2000
	Group Art Unit	
	Examiner Name	

As a below named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

Tumor-associated antigen (R11)

the specification of which (Title of the Invention)

☐ is attached hereto
OR
☒ was filed on (MM/DD/YYYY) August 3, 2000 as United States Application Number or PCT International Application Number and was amended on (MM/DD/YYYY) (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. 119(a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Priority Not Claimed	Certified Copy Attached?	
				YES	NO
DE 199 36 563.6	Germany	August 4, 1999	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

☐ Additional foreign application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto:

I hereby claim the benefit under 35 U.S.C. 119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)	
60/154,161	September 15, 1999	<input type="checkbox"/> Additional provisional application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

[Page 1 of 2]

Burden Hour Statement. This form is estimated to take 0.4 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

Please type a plus sign (+) inside this box →



PTO/SB/01 (12-97)
Approved for use through 9/30/00. OMB 0651-0032
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

DECLARATION — Utility or Design Patent Application

I hereby claim the benefit under 35 U.S.C. 120 of any United States application(s), or 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application or PCT Parent Number	Parent Filing Date (MM/DD/YYYY)	Parent Patent Number (if applicable)

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

As a named inventor, I hereby appoint the following registered practitioner(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

☒ Customer Number

000023703

OR

☐ Registered practitioner(s) name/registration number listed below

Place Customer
Number Bar Code
Label here

Name	Registration Number	Name	Registration Number
Robert P. Raymond	25,089	Susan K. Pocchiari	45,016
Alan R. Stempel	28,991	Philip I. Datlow	40,232
Marv-Ellen M. Devlin	27,928	Timothy X. Witkowski	40,232
Anthony P. Bottino	41,629		

☐ Additional registered practitioner(s) named on supplemental Registered Practitioner Information sheet PTO/SB/02C attached hereto.

Direct all correspondence to: ☒ Customer Number or Bar Code Label

000023703

OR

☐ Correspondence address below

Name	Robert P. Raymond				
Address	Boehringer Ingelheim Corporation				
Address	900 Ridgebury Road, P. O. Box 368				
City	Ridgefield	State	CT	ZIP	06877-0368
Country	US	Telephone	203/798-9988	Fax	203/798-4408

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Name of Sole or First Inventor:

☐ A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])		Family Name or Surname				
Renate		Konopitzky				
Inventor's Signature		Date				
Residence: City	Bad Voeslau	State	Country	Austria	Citizenship	AT
Post Office Address	Waldwiese 6/11					
Post Office Address						
City	Bad Voeslau	State	ZIP	A-2540	Country	AT

☒ Additional inventors are being named on the supplemental Additional Inventor(s) sheet(s) PTO/SB/02A attached hereto

Please type a plus sign (+) inside this box → +

PTO/SB/02A (3-97)
Approved for use through 9/30/98. OMB 0651-0032
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

+

DECLARATION	ADDITIONAL INVENTOR(S) Supplemental Sheet Page <u>1</u> of <u>1</u>
--------------------	---

Name of Additional Joint Inventor, if any:		<input type="checkbox"/> A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Ulrich				Koenig			
Inventor's Signature						Date	
Residence: City	Vienna	State		Country	Austria	Citizenship	AT
Post Office Address: Martinstrasse 90/2/9							
Post Office Address:							
City	Vienna	State		ZIP	A-1180	Country	Austria
Name of Additional Joint Inventor, if any:		<input type="checkbox"/> A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Wolfgang				Sommergruber			
Inventor's Signature						Date	
Residence: City	Purkersdorf	State		Country	Austria	Citizenship	AT
Post Office Address: Linzerstrasse 19/Haus 4							
Post Office Address:							
City	Purkersdorf	State		ZIP	A-3002	Country	Austria
Name of Additional Joint Inventor, if any:		<input type="checkbox"/> A petition has been filed for this unsigned inventor					
Given Name (first and middle [if any])				Family Name or Surname			
Thomas				Woelfel			
Inventor's Signature						Date	
Residence: City	Lion	State		Country	Germany	Citizenship	DE
Post Office Address: Feuchtwangerstrasse 15							
Post Office Address:							
City	Lion	State		ZIP	D-55129	Country	Germany

Burden Hour Statement: This form is estimated to take 0.4 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

+

Case No. 12/211

<110> Boehringer Ingelheim International GmbH

<130> 12211aa

<140>

<141>

<160> 102

<170> PatentIn Ver. 2.1

<210> 1

<211> 6582

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR

<222> (1) .. (217)

<220>

<221> CDS

<222> (218) .. (1423)

<220>

<221> unsure

<222> (1424) .. (1497)

$\langle 220 \rangle$

<221> CDS

<222> (1498) .. (2571)

<220>

<221> 3'UTR

<222> (2572) .. (6582)

<400> 1

acacgcgctt caacttcggt tgggtgtgtg cgaagaaacc tgactgcgcc ctgaggagaa 60

cagcggagaa ggtccacoga gcctggcgaa aggtccgctg agcgggctgt cgtccggagc 120

cactccgggc tgoggagcac ccagtggaga cggcgctgg ctcaggtgtg ggaccccatc 180

cttcctgtct tgcagagga gtccctogcgt ggtgagt atg cga aat aag cgg gtt 235
Met Arg Asn Lys Arg Val

1

5

ttg aaa aca aaa aaa aga agg agt gga aga ggg ggc cag gat cca ggc 283
Leu Lys Thr Lys Lys Arg Arg Ser Gly Arg Gly Gly Gln Asp Pro Gly

10

15

20

ctc cat ccc cac aga agt gaa gct aca gct ggg agg tct cct ccc acc 331
 Leu His Pro His Arg Ser Glu Ala Thr Ala Gly Arg Ser Pro Pro Thr
 25 30 35

cca acc gtc acc ctg ggt ccc gac tgc cca cct cct cct cct ccc cct 379
 Pro Thr Val Thr Leu Gly Pro Asp Cys Pro Pro Pro Pro Pro Pro Pro
 40 45 50

ccc ccc aac aac aac aac aac aac aac tcc aag cac acc ggc cat aag 427
 Pro Pro Asn Asn Asn Asn Asn Asn Asn Ser Lys His Thr Gly His Lys
 55 60 65 70

agt gcg tgt gtc ccc aac atg acc gaa cga aga agg gac gag ctc tct 475
 Ser Ala Cys Val Pro Asn Met Thr Glu Arg Arg Arg Asp Glu Leu Ser
 75 80 85

gaa gag atc aac aac tta aga gag aag gtc atg aag cag tcg gag gag 523
 Glu Glu Ile Asn Asn Leu Arg Glu Lys Val Met Lys Gln Ser Glu Glu
 90 95 100

aac aac aac ctg cag agc cag gtg cag aag ctc aca gag gag aac acc 571
 Asn Asn Asn Leu Gln Ser Gln Val Gln Lys Leu Thr Glu Glu Asn Thr
 105 110 115

acc ctt cga gag caa gtg gaa ccc acc cct gag gat gag gat gat gac 619
 Thr Leu Arg Glu Gln Val Glu Pro Thr Pro Glu Asp Glu Asp Asp Asp
 120 125 130

atc gag ctc cgc ggt gct gca gca gct gct gcc cca ccc cct cca ata 667
 Ile Glu Leu Arg Gly Ala Ala Ala Ala Ala Pro Pro Pro Pro Ile
 135 140 145 150

gag gaa gag tgc cca gaa gac ctc cca gag aag ttc gat ggc aac cca 715
 Glu Glu Glu Cys Pro Glu Asp Leu Pro Glu Lys Phe Asp Gly Asn Pro
 155 160 165

gac atg ctg gct cct ttc atg gcc cag tgc cag atc ttc atg gaa aag 763
 Asp Met Leu Ala Pro Phe Met Ala Gln Cys Gln Ile Phe Met Glu Lys
 170 175 180

agc acc agg gat ttc tca gtt gat cgt gtc cgt gtc tgc ttc gtg aca 811
 Ser Thr Arg Asp Phe Ser Val Asp Arg Val Arg Val Cys Phe Val Thr
 185 190 195

agc atg atg acc ggc cgt gct gcc cgt tgg gcc tca gca aag ctg gag 859
 Ser Met Met Thr Gly Arg Ala Ala Arg Trp Ala Ser Ala Lys Leu Glu
 200 205 210

cgc tcc cac tac ctg atg cac aac tac cca gct ttc atg atg gaa atg 907
 Arg Ser His Tyr Leu Met His Asn Tyr Pro Ala Phe Met Met Glu Met
 215 220 225 230

aag cat gtc ttt gaa gac cct cag agg cga gag gtt gcc aaa cgc aag 955
 Lys His Val Phe Glu Asp Pro Gln Arg Arg Glu Val Ala Lys Arg Lys
 235 240 245

atc aga cgc ctg cgc caa ggc atg ggg tct gtc atc gac tac toc aat Ile Arg Arg Leu Arg Gln Gly Met Gly Ser Val Ile Asp Tyr Ser Asn 250 255 260	1003
gct ttc cag atg att gcc cag gac ctg gat tgg aac gag cct gog ctg Ala Phe Gln Met Ile Ala Gln Asp Leu Asp Trp Asn Glu Pro Ala Leu 265 270 275	1051
att gac cag tac cac gag ggc ctc agc gac cac att cag gag gag ctc Ile Asp Gln Tyr His Glu Gly Leu Ser Asp His Ile Gln Glu Glu Leu 280 285 290	1099
toc cac ctc gag gtc gcc aag tog ctg tct gct ctg att ggg cag tgc Ser His Leu Glu Val Ala Lys Ser Leu Ser Ala Leu Ile Gly Gln Cys 295 300 305 310	1147
att cac att gag aga agg ctg gcc agg gct gct gca gct cgc aag cca Ile His Ile Glu Arg Arg Leu Ala Arg Ala Ala Ala Arg Lys Pro 315 320 325	1195
cgc tog cca ccc ogg gog ctg gtg ttg cct cac att gca agc cac cac Arg Ser Pro Pro Arg Ala Leu Val Leu Pro His Ile Ala Ser His His 330 335 340	1243
cag gta gat cca acc gag cgg gtg gga ggt gcc cgc atg cgc ctg acg Gln Val Asp Pro Thr Glu Pro Val Gly Gly Ala Arg Met Arg Leu Thr 345 350 355	1291
cag gaa gaa aaa gaa aga cgc aga aag ctg aac ctg tgc ctc tac tgt Gln Glu Glu Lys Glu Arg Arg Lys Leu Asn Leu Cys Leu Tyr Cys 360 365 370	1339
gga aca gga ggt cac tac gct gac aat tgt cct gcc aag gcc tca aag Gly Thr Gly Gly His Tyr Ala Asp Asn Cys Pro Ala Lys Ala Ser Lys 375 380 385 390	1387
tct tog cgg gog gga aac toc cgg gcc cgg ctg tag agggaccttc Ser Ser Pro Ala Gly Asn Ser Pro Ala Pro Leu 395 400	1433
agcgacgggg ccagaaataa taagggtcccc acaagatgat gctcatctc cacacttgca	1493
agtg atg ctc cag att cat ctt cog ggc aga cac acc ctg ttc gtc oga Met Leu Gln Ile His Leu Pro Gly Arg His Thr Leu Phe Val Arg 405 410 415	1542
gcc atg atc gat tct ggt gct tct ggc aac ttc att gat cac gaa tat Ala Met Ile Asp Ser Gly Ala Ser Gly Asn Phe Ile Asp His Glu Tyr 420 425 430	1590
gtt gct caa aat gga att cct cta aga atc aag gac tgg cca ata ctt Val Ala Gln Asn Gly Ile Pro Leu Arg Ile Lys Asp Trp Pro Ile Leu 435 440 445	1638
gtg gaa gca att gat ggg cgc ccc ata gca tog ggc cca gtt gtc cac Val Glu Ala Ile Asp Gly Arg Pro Ile Ala Ser Gly Pro Val Val His 450 455 460 465	1686

gaa act cac gac ctg ata gtt gac ctg gga gat cac oga gag gtg ctg 1734
 Glu Thr His Asp Leu Ile Val Asp Leu Gly Asp His Arg Glu Val Leu
 470 475 480

tca ttt gat gtg act cag tct cca ttc ttc cct gtc gtc cta ggg gtt 1782
 Ser Phe Asp Val Thr Gln Ser Pro Phe Phe Pro Val Val Leu Gly Val
 485 490 495

ogc tgg ctg agc aca cat gat ccc aat atc aca tgg agc act oga tct 1830
 Arg Trp Leu Ser Thr His Asp Pro Asn Ile Thr Trp Ser Thr Arg Ser
 500 505 510

atc gtc ttt gat tct gaa tac tgc ogc tac cac tgc ogc atg tat tct 1878
 Ile Val Phe Asp Ser Glu Tyr Cys Arg Tyr His Cys Arg Met Tyr Ser
 515 520 525

cca ata cca cca tog ctc cca cca cca gca cca caa cog cca ctc tat 1926
 Pro Ile Pro Pro Ser Leu Pro Pro Pro Ala Pro Gln Pro Pro Leu Tyr
 530 535 540 545

tat cca gta gat gga tac aga gtt tac caa cca gtg agg tat tac tat 1974
 Tyr Pro Val Asp Gly Tyr Arg Val Tyr Gln Pro Val Arg Tyr Tyr Tyr
 550 555 560

gtc cag aat gtg tac act cca gta gat gag cac gtc tac cca gat cac 2022
 Val Gln Asn Val Tyr Thr Pro Val Asp Glu His Val Tyr Pro Asp His
 565 570 575

ogc ctg gtt gac cct cac ata gaa atg ata cct gga gca cac agt att 2070
 Arg Leu Val Asp Pro His Ile Glu Met Ile Pro Gly Ala His Ser Ile
 580 585 590

ccc agt gga cat gtg tat tca ctg tcc gaa cct gaa atg gca gct ctt 2118
 Pro Ser Gly His Val Tyr Ser Leu Ser Glu Pro Glu Met Ala Ala Leu
 595 600 605

oga gat ttt gtg gca aga aat gta aaa gat ggg cta att act cca aog 2166
 Arg Asp Phe Val Ala Arg Asn Val Lys Asp Gly Leu Ile Thr Pro Thr
 610 615 620 625

att gca cct aat gga gcc caa gtt ctc cag gtg aag agg ggg tgg aaa 2214
 Ile Ala Pro Asn Gly Ala Gln Val Leu Gln Val Lys Arg Gly Trp Lys
 630 635 640

ctg caa gtt tct tat gat tgc oga gct cca aac aat ttt act atc cag 2262
 Leu Gln Val Ser Tyr Asp Cys Arg Ala Pro Asn Asn Phe Thr Ile Gln
 645 650 655

aat cag tat cct ogc cta tct att cca aat tta gaa gac caa gca cac 2310
 Asn Gln Tyr Pro Arg Leu Ser Ile Pro Asn Leu Glu Asp Gln Ala His
 660 665 670

ctg gca aog tac act gaa ttc gta cct caa ata cct gga tac caa aca 2358
 Leu Ala Thr Tyr Thr Glu Phe Val Pro Gln Ile Pro Gly Tyr Gln Thr
 675 680 685

CCGC "CGTGGG"

tac ccc aca tat gcc ggc tac ccg acc tac cca gta gga ttc gcc tgg	2406
Tyr Pro Thr Tyr Ala Ala Tyr Pro Thr Tyr Pro Val Gly Phe Ala Trp	
690 695 700 705	
tac cca gtg gga oga gac gga caa gga aga tca cta tat gta cct gtg	2454
Tyr Pro Val Gly Arg Asp Gly Gln Gly Arg Ser Leu Tyr Val Pro Val	
710 715 720	
atg atc act tgg aat cca cac tgg tac cgc cag cct ccg gta cca cag	2502
Met Ile Thr Trp Asn Pro His Trp Tyr Arg Gln Pro Pro Val Pro Gln	
725 730 735	
tac ccg ccg cca cag ccg ccg cct cca cca cca cca ccg ccg ccg cct	2550
Tyr Pro Pro Pro Gln Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro	
740 745 750	
cca tct tac agt acc ctg taa atacctgtca tgtccttcag gatctctgcc	2601
Pro Ser Tyr Ser Thr Leu	
755 760	
ctcaaaattt attcctgttc agctttctcaa tcagtgactg tgtgctaaat ttagggtac	2661
tgtatcttca ggcacactga ggcacatcct ctctgaaacg gctatggaag gttagggcca	2721
ctctggactg gcacacatcc taaagcacca aaagaccttc aacattttct gagagcaaca	2781
gagtatttgc caataaatga tctctcatit ttccaccttg actgccaatc taactaaaat	2841
aattaataag tttactttcc agccagtoct ggaagtctgg gttttacctg ccaaaacctc	2901
catcaccatc taaattatag gctgccaaat ttgctgttta acatttacag agaagctgat	2961
acaaaocgag gaaatgctga tttctttatg gagggggaga cgaggaggag gaggacatga	3021
cttttcttgc ggtttoggta cctctttttt aaatcactgg aggactgagg ccttattaag	3081
gaagccaaaa ttatcgggtgc agtgtggaaa ggcttcocgtg atcctctcgc tgcaccotta	3141
gaaacttcac cgtcttcaaa ctccatttcc atggttctgt taattctcaa ggagcagcaa	3201
ctcgactggg tctccagga gcaggaaaaa ccttgtgac atgaacatc tcaggcctga	3261
aaagaaagtg ctctctcaga tggactcttg catgttaaga ctatgtcttc acatcatggg	3321
gcaaatcaca tgtaccaat gactcoggct ttgacacaac acctaccat catcatgcca	3381
tgatggcttc cacaaagcat taaacctggg aaccagagat tactgggtggc tccagcgttg	3441
ttagatgttc atgaaatgtg accacctctc aatcaccttt gagggctaaa gagtagcaca	3501
tcaaaaggac tccaaaatcc catacccaac tcttaagaga tttgtcctgg tacttcagaa	3561
agaattttca tgagtgttct taattggctg gaaaagcacc agctgacgtt ttggaagaat	3621
ctatccatgt gtctgcctcc atatgcatct gggcatttca tcttcagtcc cctcattaga	3681
ctgtagcatt aggatgtgtg gagagaggag aatgatatta gcaccagat tcacactcct	3741

atgcctggaa gggggacatc ttgaagaag aggaattagg gctgtggaca ctgtcttgag 3801
gatgtggact tccttagtga gctccacatt acttgatggt aaccacttca aaaggatcag 3861
aatccaogta atgaaaaagg tcctctaga ggatggagct gatgtgaagc tgccaatgga 3921
tgaaaagcct cagaaaagca ctcaaaggac tcaaagcaac ggacaacaca agagttgtct 3981
tcagcccagt gacacctctg atgtcccctg gaagctttgt gctaacctgg gactgctga 4041
cttcctttag cctggtcct tgctactacc ttgaactgtt ttatctaacc tctcttttc 4101
tgtttaattc ttgtctactg ccattgacct tgctgcagga ttgtgtcat tttcctgcct 4161
ggttgctgag actccatttt gctgccacac acagagatgt aagaggcagg cttaattgc 4221
caaagcacag tttagcagc agaaaacaac atgggtgata tctcaaattg cctgacatga 4281
agaggagtct aaoggtgaag ttactcttt catcagcatc atctttcaca tgttcattat 4341
catcogctct tattcttgca tgtttaaaca cttaaaattt ttagtataat ttttagtgtg 4401
ttttgaagtg gtgactagc tttcaaaaac ttccattgaa ttacaaagca ctatccagtt 4461
cttattgtta aactaagtaa aaatgataag taacatagtg taaaatattc ctttactgtg 4521
aacttcttac aatgctgtga atgagaggct cctcagaact ggagcatttg tataataatt 4581
catcctgttc atcttcaatt ttaacatcat atataattc aattctatca attgggcctt 4641
taaaaatcat ataaaaggat ataaaattg aaaagagaaa cctaattggc tatttaatcc 4701
aaaacaactt ttttttttcc ttcaatggaa tcagaaagct tgtcaatcac tcatgtgttt 4761
tagagtaatt acttttaaaa tgggtgcattt gtgcttctga actattttga agagtcactt 4821
ctgtttacct caagtatcaa ttcatcctcc atacatttga attcaagttg ttttttgtca 4881
aatttacagt tgtcaattga tcttcaagct gcagggtgcc tagaaatggg cagtgtctg 4941
tagccctggc atgtgcacac ggacatttgc caccactgca agcaaaagtc tggagaagtt 5001
caccaaogac aagaacgatt agggaaaata tgctgctgtg ggttaacaac tcagaaagtc 5061
cctgatccac atttggctgt ttactaaagc ttgtgattaa ctttttgga gtgtgtacta 5121
tgctctattg ctatatatgc tatctataaa tgtagatgtt aaggataagt aattctaaat 5181
ttattattct atagttttga agtttggtta agtttccttt cactcaattg atttattttg 5241
ttgttaatca aatttatgtt aattggatcc tttaaatttt ttttggcatt ttccaacaaa 5301
aatggcttta ttcataagaa agggaaaaaa tcaatggaat ttgatatcta aagaagttag 5361
aaaggagca aaataaaaaa cataaaggag atagatgaat tagtaagcaa atcagtagtc 5421
gagtttttca aactggcaaa attaattaat tgacttttag cccaatttta cattgttaat 5481

taaatcaaga aggaagaaga tctaagagct cccattgata ggcaagoccta gagagaacta 5541
 gctaaattta tcatgctagg atattgaaac acagaaagtt tacatacatt tatgaagggg 5601
 caatttagtt tggacagtga ggtatttgtc ttagtggaaa aaaggagaat tagtctgac 5661
 aaatogtgaa gtaatacagt gaacttgacg gtgcacaaaa taagagggcc acatctatat 5721
 ggtgcagtct ggaattctgt ttaagtttgt aggtacctct tggacttctg aattgatcca 5781
 gttgtcatcc accacagaca tctcacatca gatacagaca gttccaagat tgacaacaga 5841
 gaacaacctg ctggaaagac ctgggcagaa atggagagcc ctgogggaac catgctacat 5901
 tttcatctaa agagagaatg cacatctgat gagactgaaa gttctttggt gtttagatt 5961
 gtagaatggt attgaattgg tctgtggaaa attgcattgc ttttatttct ttgtgtaac 6021
 aagtttaagt aataggggat atataatcat aagcatttta ggggtggagg gactattaag 6081
 taattttaag tgggtgggt tatttagaat gttagaataa tattatgtat tagatatgc 6141
 tataagtga catgogtact tacttgaac cctttacct ataattgcta tccttaaaga 6201
 tttcaaataa actcggagg aactgcagg agaccaactt atttagagcg aattggacat 6261
 ggataaaaac cccagtggga gaaagttcaa aggtgattag attaataatt taatagagga 6321
 tgagtgcct ctgataaatt actgctagaa tgaacttgc aatgatggat ggtaaatttt 6381
 catggaagtt ataaaagtga taaataaaaa ccttgcttt taccctgtc agtagccctc 6441
 ctctaccac tgaacccat tgccctacc cctcttcta actttattgc tgtattctct 6501
 tcactctata tttctctcta ttgctaata ttgcattgct gttacaataa aaattcaata 6561
 aagatttagt ggttaagtgc t 6582

<210> 2

<211> 401

<212> PRT

<213> Homo sapiens

<400> 2

Met	Arg	Asn	Lys	Arg	Val	Leu	Lys	Thr	Lys	Lys	Arg	Arg	Ser	Gly	Arg
1					5				10					15	
Gly	Gly	Gln	Asp	Pro	Gly	Leu	His	Pro	His	Arg	Ser	Glu	Ala	Thr	Ala
			20					25					30		
Gly	Arg	Ser	Pro	Pro	Thr	Pro	Thr	Val	Thr	Leu	Gly	Pro	Asp	Cys	Pro
		35					40				45				
Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Asn	Asn	Asn	Asn	Asn	Asn	Asn	Ser
		50					55				60				

Lys His Thr Gly His Lys Ser Ala Cys Val Pro Asn Met Thr Glu Arg
 65 70 75 80
 Arg Arg Asp Glu Leu Ser Glu Glu Ile Asn Asn Leu Arg Glu Lys Val
 85 90 95
 Met Lys Gln Ser Glu Glu Asn Asn Asn Leu Gln Ser Gln Val Gln Lys
 100 105 110
 Leu Thr Glu Glu Asn Thr Thr Leu Arg Glu Gln Val Glu Pro Thr Pro
 115 120 125
 Glu Asp Glu Asp Asp Asp Ile Glu Leu Arg Gly Ala Ala Ala Ala Ala
 130 135 140
 Ala Pro Pro Pro Pro Ile Glu Glu Glu Cys Pro Glu Asp Leu Pro Glu
 145 150 155 160
 Lys Phe Asp Gly Asn Pro Asp Met Leu Ala Pro Phe Met Ala Gln Cys
 165 170 175
 Gln Ile Phe Met Glu Lys Ser Thr Arg Asp Phe Ser Val Asp Arg Val
 180 185 190
 Arg Val Cys Phe Val Thr Ser Met Met Thr Gly Arg Ala Ala Arg Trp
 195 200 205
 Ala Ser Ala Lys Leu Glu Arg Ser His Tyr Leu Met His Asn Tyr Pro
 210 215 220
 Ala Phe Met Met Glu Met Lys His Val Phe Glu Asp Pro Gln Arg Arg
 225 230 235 240
 Glu Val Ala Lys Arg Lys Ile Arg Arg Leu Arg Gln Gly Met Gly Ser
 245 250 255
 Val Ile Asp Tyr Ser Asn Ala Phe Gln Met Ile Ala Gln Asp Leu Asp
 260 265 270
 Trp Asn Glu Pro Ala Leu Ile Asp Gln Tyr His Glu Gly Leu Ser Asp
 275 280 285
 His Ile Gln Glu Glu Leu Ser His Leu Glu Val Ala Lys Ser Leu Ser
 290 295 300
 Ala Leu Ile Gly Gln Cys Ile His Ile Glu Arg Arg Leu Ala Arg Ala
 305 310 315 320
 Ala Ala Ala Arg Lys Pro Arg Ser Pro Pro Arg Ala Leu Val Leu Pro
 325 330 335
 His Ile Ala Ser His His Gln Val Asp Pro Thr Glu Pro Val Gly Gly
 340 345 350
 Ala Arg Met Arg Leu Thr Gln Glu Glu Lys Glu Arg Arg Arg Lys Leu
 355 360 365

Asn Leu Cys Leu Tyr Cys Gly Thr Gly Gly His Tyr Ala Asp Asn Cys
 370 375 380

Pro Ala Lys Ala Ser Lys Ser Ser Pro Ala Gly Asn Ser Pro Ala Pro
 385 390 395 400

Leu

<210> 3

<211> 357

<212> PRT

<213> Homo sapiens

<400> 3

Met Leu Gln Ile His Leu Pro Gly Arg His Thr Leu Phe Val Arg Ala
 1 5 10 15

Met Ile Asp Ser Gly Ala Ser Gly Asn Phe Ile Asp His Glu Tyr Val
 20 25 30

Ala Gln Asn Gly Ile Pro Leu Arg Ile Lys Asp Trp Pro Ile Leu Val
 35 40 45

Glu Ala Ile Asp Gly Arg Pro Ile Ala Ser Gly Pro Val Val His Glu
 50 55 60

Thr His Asp Leu Ile Val Asp Leu Gly Asp His Arg Glu Val Leu Ser
 65 70 75 80

Phe Asp Val Thr Gln Ser Pro Phe Phe Pro Val Val Leu Gly Val Arg
 85 90 95

Trp Leu Ser Thr His Asp Pro Asn Ile Thr Trp Ser Thr Arg Ser Ile
 100 105 110

Val Phe Asp Ser Glu Tyr Cys Arg Tyr His Cys Arg Met Tyr Ser Pro
 115 120 125

Ile Pro Pro Ser Leu Pro Pro Pro Ala Pro Gln Pro Pro Leu Tyr Tyr
 130 135 140

Pro Val Asp Gly Tyr Arg Val Tyr Gln Pro Val Arg Tyr Tyr Tyr Val
 145 150 155 160

Gln Asn Val Tyr Thr Pro Val Asp Glu His Val Tyr Pro Asp His Arg
 165 170 175

Leu Val Asp Pro His Ile Glu Met Ile Pro Gly Ala His Ser Ile Pro
 180 185 190

Ser Gly His Val Tyr Ser Leu Ser Glu Pro Glu Met Ala Ala Leu Arg
 195 200 205

Asp Phe Val Ala Arg Asn Val Lys Asp Gly Leu Ile Thr Pro Thr Ile
 210 215 220

CCPDB v. 2019.04

Ala Pro Asn Gly Ala Gln Val Leu Gln Val Lys Arg Gly Trp Lys Leu
 225 230 235 240

Gln Val Ser Tyr Asp Cys Arg Ala Pro Asn Asn Phe Thr Ile Gln Asn
 245 250 255

Gln Tyr Pro Arg Leu Ser Ile Pro Asn Leu Glu Asp Gln Ala His Leu
 260 265 270

Ala Thr Tyr Thr Glu Phe Val Pro Gln Ile Pro Gly Tyr Gln Thr Tyr
 275 280 285

Pro Thr Tyr Ala Ala Tyr Pro Thr Tyr Pro Val Gly Phe Ala Trp Tyr
 290 295 300

Pro Val Gly Arg Asp Gly Gln Gly Arg Ser Leu Tyr Val Pro Val Met
 305 310 315 320

Ile Thr Trp Asn Pro His Trp Tyr Arg Gln Pro Pro Val Pro Gln Tyr
 325 330 335

Pro Pro Pro Gln Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro Pro
 340 345 350

Ser Tyr Ser Thr Leu
 355

<210> 4
 <211> 31
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 4
 ctagcccacc atggcatctg cagecacgtg a

31

<210> 5
 <211> 30
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 5
 agcttcacgt ggctgcagat gccatggtgg

30

<210> 6
 <211> 30
 <212> DNA
 <213> Synthetic Sequence

 <220>
 <223> Description of the Synthetic Sequence: Primer

 <400> 6
 ctagcccacc atggcatctg cagcacgtga 30

<210> 7
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

 <220>
 <223> Description of the Synthetic Sequence: Primer

 <400> 7
 agcttcacgt ggtgcagatg ccatggtgg 29

<210> 8
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

 <220>
 <223> Description of the Synthetic Sequence: Primer

 <400> 8
 ctagcccacc atggcatctg cacacgtga 29

<210> 9
 <211> 28
 <212> DNA
 <213> Synthetic Sequence

 <220>
 <223> Description of the Synthetic Sequence: Primer

 <400> 9
 agcttcacgt gtgcagatgc catggtgg 28

<210> 10
 <211> 23
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 10
 gggcggtagg cgtgtacggt ggg 23

<210> 11
 <211> 26
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 11
 gcaactagaa ggcacagtcg aggctg 26

<210> 12
 <211> 27
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 12
 gtttggacag tgaggtattt gtcttag 27

<210> 13
 <211> 27
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 13
 ctttcagca ggttggtctc tgtgtgc 27

<210> 14
 <211> 30
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: synthetic Primer

<400> 14
 tgacgggggtc acccacactg tgcccatcta

30

<210> 15
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 15
 ctagaagcat tgcggtggac gatggaggg

29

<210> 16
 <211> 22
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 16
 aaggtgaagg tcggagtcaa cg

22

<210> 17
 <211> 24
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 17
 ggcagagatg atgacccttt tggc

24

<210> 18
 <211> 26
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 18
 tattttgctc cctttotaac ttcttt

26

<210> 19
 <211> 30
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 19
 ttctactttt catcagcatc atctttcaca

30

<210> 20
 <211> 27
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 20
 cgttagactc ctcttcatgt caggcaa

27

<210> 21
 <211> 22
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 21
 ggtgacacta tagaaggtag gc

22

<210> 22
 <211> 28
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 22
 caggcctgag atgtttcatg tcacaagg

28

<210> 23
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 23
 gcatttcctg cgtttgtatc agcttctct

29

<210> 24
 <211> 30
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 24
 accagcacca caacggccac tctattatcc

30

<210> 25
 <211> 30
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 25
 catatagtga tcttccttgt cgtctcgtc

30

<210> 26
 <211> 26
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 26
 gggcccatca attgcttcca caagta

26

<210> 27
 <211> 21
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 27
 gcagagctcg tttagtgaac c

21

<210> 28
 <211> 30
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 28
 ggccagaaat aataaggtcc ccacaagatg

30

<210> 29
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

<220>
 <223> Description of the Synthetic Sequence: Primer

<400> 29
 agctttctgc gtctttcttt ttcttctg

29

<213> Synthetic Sequence

<223> Description of the Synthetic Sequence: Primer

26

<213> Synthetic Sequence

<223> Description of the Synthetic Sequence: Primer

37

<213> Synthetic Sequence

<223> Description of the Synthetic Sequence: Primer

27

<213> Synthetic Sequence

<223> Description of the Synthetic Sequence: Primer

22

<400> 37
cagggtagcg gttggggtagg gaggagac 28

<210> 38
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 38
 gcttcacttc tgtggggatg gaggcctgg

29

<210> 39
 <211> 22
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 39
 atgcgaaata agcgggtttt ga

22

<210> 40
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 40
 cgcagaggag tcctcgcgtg gtgagtatg

29

<210> 41
 <211> 29
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 41
 ggctcaggtg tgggacccca tccttcctg

29

<210> 42

<211> 29
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 42
 gctccggaag acagcccgt cagcggacc

29

<210> 43
 <211> 24
 <212> DNA
 <213> Synthetic Sequence

<220>

<223> Description of the Synthetic Sequence: Primer

<400> 43
 gaagaaacct gactgcgcc tgag

24

<210> 44
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 44
 Met Leu Gln Ile His Leu Pro Gly Arg
 1 5

<210> 45
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 45
 Ile His Leu Pro Gly Arg His Thr Leu
 1 5

<210> 46
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 46

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

His Leu Pro Gly Arg His Thr Leu Phe
1 5

<210> 47
<211> 9
<212> PRT
<213> Homo sapiens

<400> 47
Tyr Val Ala Gln Asn Gly Ile Pro Leu
1 5

<210> 48
<211> 9
<212> PRT
<213> Homo sapiens

<400> 48
Leu Arg Ile Lys Asp Trp Pro Ile Leu
1 5

<210> 49
<211> 9
<212> PRT
<213> Homo sapiens

<400> 49
Ile Leu Val Glu Ala Ile Asp Gly Arg
1 5

<210> 50
<211> 9
<212> PRT
<213> Homo sapiens

<400> 50
Gly Arg Pro Ile Ala Ser Gly Pro Val
1 5

<210> 51
<211> 9
<212> PRT
<213> Homo sapiens

CCDS: C12345.1


```
<210> 61
<211> 9
<212> PRT
<213> Homo sapiens
```

<400> 61
 Arg Val Tyr Gln Pro Val Arg Tyr Tyr
 1 5

<210> 62
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 62
 Tyr Gln Pro Val Arg Tyr Tyr Tyr Val
 1 5

<210> 63
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 63
 Val Arg Tyr Tyr Tyr Val Gln Asn Val
 1 5

<210> 64
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 64
 Tyr Val Gln Asn Val Tyr Thr Pro Val
 1 5

<210> 65
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 65
 Glu His Val Tyr Pro Asp His Arg Leu
 1 5

<210> 66
 <211> 9
 <212> PRT
 <213> Homo sapiens

CCDS: C10000.1

<400> 66
 Leu Val Asp Pro His Ile Glu Met Ile
 1 5

<210> 67
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 67
 Glu Met Ile Pro Gly Ala His Ser Ile
 1 5

<210> 68
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 68
 His Ser Ile Pro Ser Gly His Val Tyr
 1 5

<210> 69
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 69
 Ile Pro Ser Gly His Val Tyr Ser Leu
 1 5

<210> 70
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 70
 Ser Leu Ser Glu Pro Glu Met Ala Ala
 1 5

<210> 71
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 71
 Leu Ser Glu Pro Glu Met Ala Ala Leu
 1 5

<210> 72
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 72
 Pro Glu Met Ala Ala Leu Arg Asp Phe
 1 5

<210> 73
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 73
 Glu Met Ala Ala Leu Arg Asp Phe Val
 1 5

<210> 74
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 74
 Ala Leu Arg Asp Phe Val Ala Arg Asn
 1 5

<210> 75
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 75
 Val Ala Arg Asn Val Lys Asp Gly Leu
 1 5

<210> 76
 <211> 9
 <212> PRT
 <213> Homo sapiens

CCDS: C03760

```
<210> 77
<211> 9
<212> PRT
<213> Homo sapiens
```

```
<210> 78
<211> 9
<212> PRT
<213> Homo sapiens
```

```
<210> 79
<211> 9
<212> PRT
<213> Homo sapiens
```

```
<210> 80
<211> 9
<212> PRT
<213> Homo sapiens
```

```
<210> 81
<211> 9
<212> PRT
<213> Homo sapiens
```

<400> 81
 Glu Asp Gln Ala His Leu Ala Thr Tyr
 1 5

<210> 82
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 82
 His Leu Ala Thr Tyr Thr Glu Phe Val
 1 5

<210> 83
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 83
 Gly Arg Asp Gly Gln Gly Arg Ser Leu
 1 5

<210> 84
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 84
 Arg Asp Gly Gln Gly Arg Ser Leu Tyr
 1 5

<210> 85
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 85
 Asp Gly Gln Gly Arg Ser Leu Tyr Val
 1 5

<210> 86
 <211> 9
 <212> PRT
 <213> Homo sapiens

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

<400> 86

Ser Leu Tyr Val Pro Val Met Ile Thr
 1 5

<210> 87

<211> 9

<212> PRT

<213> Homo sapiens

<400> 87

Ile Thr Trp Asn Pro His Trp Tyr Arg
 1 5

<210> 88

<211> 9

<212> PRT

<213> Homo sapiens

<400> 88

Ser Pro Pro Thr Pro Thr Val Thr Leu
 1 5

<210> 89

<211> 9

<212> PRT

<213> Homo sapiens

<400> 89

Leu Ser Glu Glu Ile Asn Asn Leu Arg
 1 5

<210> 90

<211> 9

<212> PRT

<213> Homo sapiens

<400> 90

Lys Leu Thr Glu Glu Asn Thr Thr Leu
 1 5

<210> 91

<211> 9

<212> PRT

<213> Homo sapiens

<400> 91

Leu Thr Glu Glu Asn Thr Thr Leu Arg
 1 5

<210> 92
<211> 9
<212> PRT
<213> Homo sapiens

<400> 92
Ile Glu Leu Arg Gly Ala Ala Ala Ala
1 5

<210> 93
<211> 9
<212> PRT
<213> Homo sapiens

<400> 93
Phe Met Ala Gln Cys Gln Ile Phe Met
1 5

<210> 94
<211> 9
<212> PRT
<213> Homo sapiens

<400> 94
Ser Met Met Thr Gly Arg Ala Ala Arg
1 5

<210> 95
<211> 9
<212> PRT
<213> Homo sapiens

<400> 95
Ala Ala Arg Trp Ala Ser Ala Lys Leu
1 5

<210> 96
<211> 8
<212> PRT
<213> Homo sapiens

<400> 96
Ala Lys Leu Glu Arg Ser His Tyr Leu
1 5

<210> 97

<211> 9
 <212> PRT
 <213> Homo sapiens

<400> 97
 Gln Gly Met Gly Ser Val Ile Asp Tyr
 1 5

<210> 98
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 98
 Asn Glu Pro Ala Leu Ile Asp Gln Tyr
 1 5

<210> 99
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 99
 Arg Arg Leu Ala Arg Ala Ala Ala Ala
 1 5

<210> 100
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 100
 Lys Pro Arg Ser Pro Pro Arg Ala Leu
 1 5

<210> 101
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 101
 Arg Met Arg Leu Thr Gln Glu Glu Lys
 1 5

<210> 102
 <211> 9
 <212> PRT
 <213> Homo sapiens

<400> 102

Pro Thr Glu Pro Val Gly Gly Ala Arg
1 5